DOCTOR OF PHILOSOPHY

Exploring the molecular basis of host specific dynamic adaptation in Phytophthora capsici

Cummins, Michael

Award date:
2021

Exploring the molecular basis of host specific dynamic adaptation in *Phytophthora capsici*

Michael Cummins

This thesis is submitted for the degree of Doctor of Philosophy (PhD) to the University of Dundee

October 2021

Contents

List of Figures

List of Tables

List of Abbreviations

ABC transporter	ATP-binding cassette transporter
AMT	Agrobacterium Mediated Transformations
ATP	Adenosine triphosphate
AVR	Avirulence Genes
BLASTp/n	Basic Local Alignment Search Tool (protein/nucleotide)
BP	Base Pairs
Cas Protein	CRISPR-associated Protein
CB	Cucumber Broth
cDNA	Complementary DNA
CDS	Coding Sequence
CE	Cucumber Extract
CLF	Cucumber Leaf
CPM	Counts per Million Reads
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRN	CRinkling and Necrosis Effector
CSW	Cucumber Leaf Sandwich
CYP	Cytochrome
DE	Differential Expression
DEA	Dynamic Extreme Aneuploidy
DEG	Differentially expressed gene
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
eGFP/GFP	(Enhanced) Green Fluorescent Protein
ETI	Effector-Triggered Immunity
ETS	Effector-Triggered Suppression
FOL	*Fusarium oxysporum f. sp. Lycopersici*
G418	Geneticin
GC	Germinating Cysts
gDNA	Genomic DNA
GO term	Gene Ontology Term
GPCR	G-protein Coupled Receptors
GS	Glucosinolates
GSTs	Glutathione S-transferases
GTEN-DM	Glycerol-Tris-EDTA-NaCl Media with Detergent Mix
GTEN-T	Glycerol-Tris-EDTA-NaCl Media with Tween20
HA-tag	Hemagglutinin Tag
HDV ribozyme	Hepatitis Delta Virus Ribozyme
HH ribozyme	Hammerhead Ribozyme
HPI	Hours Post Inoculation
HR	Hypersensitivity Response
HRP	Horseradish Peroxidase
hSpCas9	Human Codon Optimised *Streptococcus pyogenes* Cas9 Protein
IM	Agrobacterium Induction Medium
INF1	Infestin 1
IVT	*in vitro* Transcription
KOG annotation	EuKaryotic Orthologous Groups annotation
LB	Lysogeny Broth
LFQ intensity	Label-free Quantitation Intensity
MAP kinase	Mitogen-activated Protein Kinase

MCS	Multiple Cloning Site
MES buffer	2-(N-morpholino)ethanesulfonic acid buffer
MFS	Major Facilitator Superfamily
mRNA	Messenger RNA
MTR	Methyltransferase
NB-LRR	Nucleotide-Binding domain and a Leucine-Rich Repeat domain receptors
NEP1	Necrosis and ethylene-inducing peptide 1
NHEJ	Non-homologous End Joining
NMRA	Nitrogen Metabolite Repression Regulator
NPB	Nutrient Pea Broth
NPT II	Neomycin Phosphotransferase II
NT	Nucleotides
OD	Optical Density
ORP1	Oxathiapiprolin Resistance Gene
ORX1	Oxidoreductase gene
PAM	Protospacer Adjacent Motif
PAMP	Pathogen Associated Molecular Patterns
PB	Pea Broth
PCA	Principal Component Analysis
PcHmp1	*Phytophthora capsici* haustorium-specific membrane protein 1
PcMuORP	*Phytophthora capsici* Mutated Oxathiapiprolin resistance gene1
PcNMRA1	*Phytophthora capsici* Nitrogen Metabolite Repression Regulator 1
PcNpp1	*Phytophthora capsici* necrosis-inducing *Phytophthora* protein 1
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
PenSeq	Pathogen Enrichment Sequencing
PMSF	Phenylmethylsulfonyl Fluoride
POD	Peroxidase
PPIase	Peptidyl-prolyl cis-trans Isomerase
PRR	Pattern Recognition Receptors
(Ps)NLS	*(P. sojea) N*uclear Localisation Signal
PTI	PAMP Triggered Immunity
PVPP	Polyvinylpolypyrrolidone
qRT-PCR.	Real-Time Quantitative Reverse Transcription PCR
R-gene	Resistance Gene
RIN	RNA Integrity Numbers
RMM	Ribosomal Maintenance Media
RNA	Ribonucleic acid
RNAi	RNA Interference
RNP	Ribonucleoprotein
ROS	Reactive Oxygen Species
RP	Ribosomal Protein
RPL	Large Ribosomal Protein Subunits
RPM	Revolutions Per Minute
rRNA	Ribosomal RNA
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RXLR	Conserved Arg–X–Leu–Arg Motif Effector
SD	Standard Deviation
SDS-page	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SEM	Standard Error of the Mean
sgRNA	Single Guide RNA
SNV	Single Nucleotide Variation
TB	Tomato Broth
TBE buffer	Tris-Borate-EDTA Buffer

T-DNA LB/RB	Transfer DNA Left and Right boarders
TE	Tomato Extract
TLF	Tomato Leaf
TMM	Weighted Trimmed Mean of M-values
TPM	Transcripts per Kilobase Million
TRAP	Translating Ribosome Affinity Purification
TRAP-SEQ	TRAP RNA Sequencing
tRNA	Transfer RNA
TSW	Tomato leaf sandwich
UV	Ultra Violet
WPT	Weeks Post Transformation
WT	Wild Type Strain

Acknowledgements

Firstly, I need to thank my supervisor Edgar Huitema. He has allowed me to explore and discover, all while providing guidance, insight and the occasional titbits of wisdom. He was always supportive and encouraging while guiding me through my PhD.

I also need to thank all the members of the Huitema lab past and present, Jasmin whose expertise in all realms of *Phytophthora* and molecular research was crucial and always generously shared. Tiago, Natalia and Victor who guided me whilst I was learning the ropes. And Rory and Nisha who were great lab mates throughout my whole PhD. I would also like to thank the multitude of visiting student, many of whose projects I helped supervise, and who always kept the lab fun and interesting. Anouk, Almedra, Feiji, Rudd, Anouk, Brian, Milou, Ross. And epically Joram Westera, who was a major help pursing the implementation of CRISPR/Cas9 in *P .capsici*.

I would also like to thank the many members of the Division of Plant Sciences, and The James Hutton Institute, the lab managers, especially Sandie Gray, and all members of technical and support staff. I would also like to thank the members of the University of Dundee Fingerprints Proteomics Facility, especially Cara Rogers for her help with analysis.

I need to thank the members of my Thesis committee, Prof. John Brown and Prof. Kevin Read for their advice and feedback throughout my PhD.

And last but certainly not least I would also like to thank the many friends from Life Sciences that made my time in Dundee super enjoyable; my Mother and Sister for their constant encouragement; and Julie for always being there and for her patience whilst waiting for me to finish my PhD.

Declaration

Candidate

The candidate is the author of the thesis. Unless otherwise stated, all references cited have been consulted by the candidate. The work of which the thesis is a record has been done by the candidate and has not been previously accepted for a higher degree

……..

Michael Cummins

Supervisor

The conditions of the relevant ordinances and regulations have been fulfilled

..…….

Dr. Edgar Huitema

Summary

Some of the worlds most devastating plant pathogens are those with broad host ranges (Kamoun et al., 2015, Dean et al., 2012). Yet there has been little investigation into the molecular mechanisms that enable broad host range (Mbengue et al., 2016). *Phytophthora capsici* is an oomycete plant pathogen with a broad host range and is well placed to become a model organism for the investigation of the mechanisms of host range and dynamic adaption in oomycete plant pathogens (Lamour et al., 2012a).

Thus, this projects aims were two fold, 1) to develop tools for the examination of *P. capsici* biology and 2) to investigate the mechanisms of host specific dynamic adaption during infection. It was decided that this project would take an unbiased, “omics” based approach in defining the dynamic adaption to multiple host plants.

Initial efforts were focused on developing a technology for the isolation of *P. capsici* translating mRNA from infected plant tissue, to avoid excessive plant material contaminating omics analysis (Chapter 2). However, while the isolation of mRNA in ribonucleic complexes was successfully established *in vitro,* implementation of this technology was not feasible in samples from the infection cycle. Based on these results, a simpler methodology was developed. To investigate early changes in gene expression, induced by host-derived signals, a method that relies on the incubation of germinating cysts with plant extracts was developed and validated. The new system was then used to carry out two separate omics experiments. First RNA-sequencing of total RNA comparing tomato extract and cucumber extract at 2, 4 and 8 hours post inoculation was conducted (Chapter 3). Followed by an identical proteomics experiment (Chapter 4). Results of these experiments point to key elements and patterns of the host specific dynamic adaptation of *P. capsici.* We have found that host

extracts were able to induce differential expression of genes, and proteins, amongst these oxidoredctase activity, and transporter proteins were found in high abundance, suggesting a necessary host-specific detoxification element to dynamic host adaptation. Proteins with phosphorylation activity, and other potential signalling molecules were also found in abundance in our suite of differentially regulated elements. However, what mechanisms induce this differential expression event, be they direct host perception or perhaps nutrient sensing, or some other indirect mechanism, is still unclear. In addition, how biologically relevant the dynamic adaptation event to host extract is to infection and colonisation of a host plant is still an open question. It was then the aim of this project to develop tools for the characterisation of the key genes identified in these two experiments. Two transformation methodologies for *P. capsici* were optimised, mainly for the use in the CRISPR/Cas9-system (Chapter 5). We were unable in this study to show the utility of the CRISPR/Cas9-system in *P. capsici*, and indeed it is still unclear how functional CRISPR/Cas will be in *Phytophthora* species in general. Overall this project was able to shed some light mechanism of host-specific dynamic adaption and did develop some tools for the future study of *P. capsici.*

Chapter 1: General Introduction

1.1 Introduction

The capacity to feed oneself and one's family, and the quality of that food is central to human health and wellbeing. However, global food security has always been and will always be a fundamental and ubiquitous concern. When considering the five major food crops, on average, 22.5% of global yield is lost to pests or pathogens each year (Savary et al., 2019). This, in addition to increasing populations and the climate crisis, leaves food stocks worryingly vulnerable (Schnitter and Berry, 2019). Thus, addressing one of the main limiting factors of food stock levels, that being crop loss due to pest and pathogen, is key to maintaining global human health and wellbeing.

Those pathogens with high economic impact, in general, share one of two factors. 1) They affect one of the major global food crops, for example *Magnaporthe oryzae,* or rice blast fungus, which endangers the main calorie source of over half the population of the world (Dean et al., 2012) or *Phytophthora infestans* which causes potato late blight and is infamous for being the main biotic cause of the Irish Potato Famine (Kamoun et al., 2015). 2) They have a broad host range, for example *Botrytis cinerea,* also known as grey mould, which boasts over 200 know host plant species (Dean et al., 2012), or the oomycete *Phytophthora ramorum* which has over 109 host plant species, and globally is the main cause of sudden oak death (Grunwald et al., 2008). While broad host ranged pathogens are widely recognised as highly impactful in term of crop loss, the majority of plant-pathogen interaction studies have focused on single pathogen-host associations. Understanding how broad host range pathogens are able to maintain virulence in multiple host plants from distinct plant lineages is an important

area for future study (Dong et al., 2015) (Mbengue et al., 2016). However, as of now, the factors that determine whether a plant is a host or non-host, and the factors that define a pathogen's host range, be it a broad host ranged or a specialist host range, are yet elusive. To alleviate the burden caused by crop loss, engineer solutions, and inform breeding programmes, it is necessary to understand the strategies employed by crop pathogens and the determinants of pathogenicity.

1.2 Coevolution of the Plant immune systems and Pathogenicity strategies

For a microbe to become a pathogen it has to overcome the innate immunity of a chosen host, pathogenesis is a specialised lifestyle that only a few microbes achieve, and only on particular plants. The plant - pathogen community exists in a state of constant flux, under constantly changing evolutionary pressures. Both plant and pathogen are in a continuous evolutionary arms race for improved immunity/pathogenicity, reciprocal adaptations in order to maintain the status quo in terms of virulence. As stated by the Red Queen to Alice, "In this place it takes all the running you can do to keep in the same place". A broad host range pathogen, presumably, must keep evolving on multiple fronts.

1.2.1 Pathogen triggered immunity, Effectors and Effector triggered immunity

Plants do not have mobile, specialized immune cells as in animals, a plant is able to mount an immune response in any tissue upon pathogen perception. That initial recognition of the pathogen by the plant by what are termed pattern recognition receptors (PRR) which recognise pathogen associated molecular patterns (PAMP) leads to PAMP triggered immunity (PTI) (Figure 1.1). This is usually a reaction to initial

contact with a pathogen and takes place on the surface of the plant and pathogen cells. PAMP are typically generic components of pathogens such as the flagellin of many bacteria or chitin of the cell walls of many fungal pathogens. For a microbe to successfully cause disease in a plant, it must evade or suppress PTI. Commonly the deployment of a suite effector molecules disables immunity and enhances susceptibility, enabling the microbe to become pathogenic and colonise the plant.

Many pathogens, including *Phytophthora* species, secrete both apoplastic effectors and intracellular cytoplasmic effectors through haustoria or similar specialised structures such as the type 3 secretion system found in bacteria. Oomycete cytoplasmic effectors are characterised by a highly conserved Arg–X–Leu–Arg (RXLR) motif, which is required for translation into the plant cytoplasm (Whisson et al., 2007). *P. capsici* has 515 of these predicted RXLR effector proteins, a similar quantity to the >500 predicted in *P. infestans*. (Haas et al., 2009). While functions of RXLRs are diverse, they commonly target processes that suppress immunity or facilitate disease progression. For example, AVR3a suppress the cell death response induced by the *P. infestans* elicitin infestin 1 (INF1), by stabilizing the host E3 ligase CMPG1 (Bos et al., 2006, Bos et al., 2010). Whereas *P. capsici* PcAvr3a12 targets and inhibits the host plant peptidyl-prolyl cis-trans isomerase (PPIase) which is a key regulator of immunity-associated functions (Fan et al., 2018). However, the majority of RXLRs identified in *Phytophthora*, remain uncharacterized. A second conserved family of effectors found in Phytophthora species are the crinklers (CRN) named for the CRinkling and Necrosis phenotype cause in plant leaves when these proteins are overexpressed. They are characterised by an N-terminal LXLFLAK motif, an N terminal signal peptide enabling translocation inside the plant cell, and a diverse C

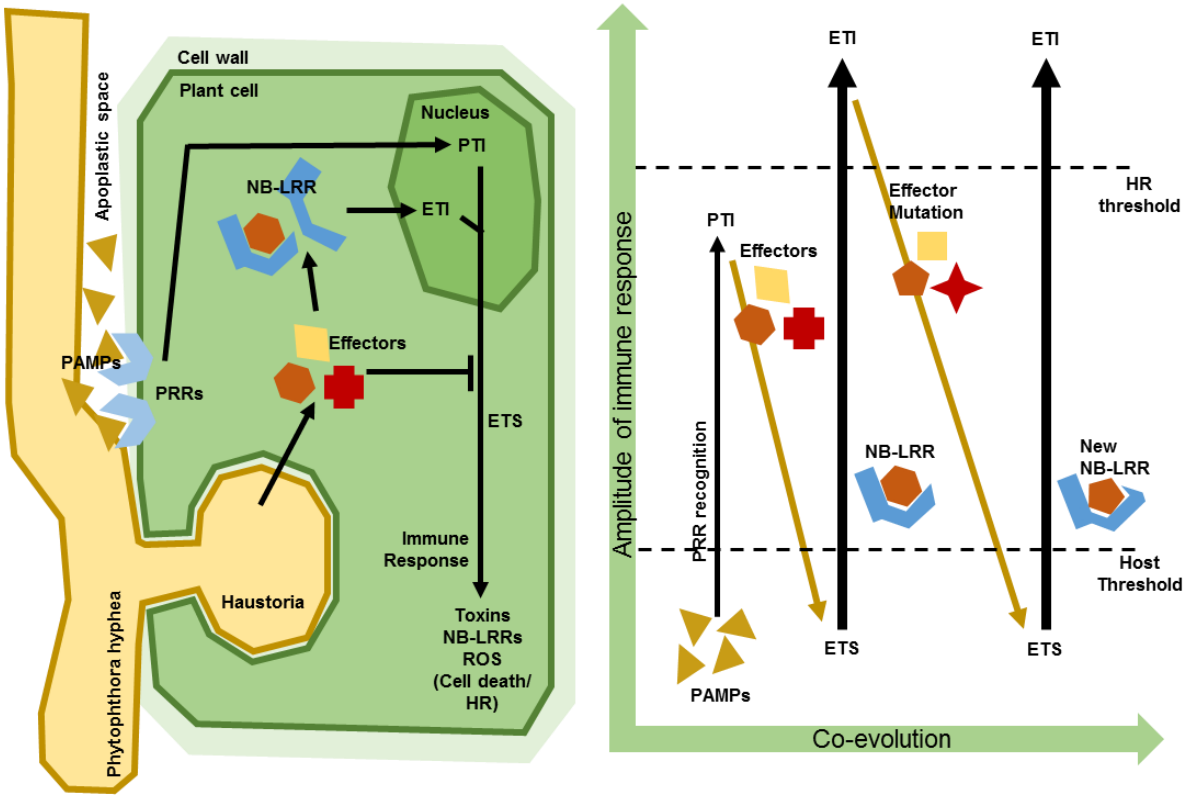


Figure 1.1: Diagram of Plant Immunity and the Zig-zag model.

Upon confrontation with a microbe and potential pathogen, initial recognition of the microbe is through conserved pathogen associated molecular patterns (PAMPS). These are recognised by pattern recognition receptors (PRR), which activates the first level of plant immunity, PAMP trigger immunity (PTI). This immune response can include toxin metabolisation, NB-LRR generation, and ROS production. For a microbe to successfully become a pathogen it has to inhibit this initial PTI. It does this through the secretion of effector proteins from specialised structures such haustoria. These effectors inhibit the immune response and allow pathogenicity, this is termed effector trigger suitability (ETS). Through years of coevolution of host and pathogen, plants have evolved a means to recognise these effectors using R-genes or NB-LRR proteins, this triggers a secondary immune response, effector triggered immunity (ETI). Along with the other hallmarks of the immune response ETI can also lead to a hypersensitivity response (HR) which is characterised by plant cell death to inhibit infection spread. Continued coevolution of host and pathogen leads to a back and forth, an arms race, where the pathogen evolves new effectors in order to more effectively inhibit the immune response and allow pathogenicity, making the plant a host. In response the plant produces new R –genes for effector recognition in order to maintain and increase the amplitude of the immune response and obstruct the pathogenicity of the microbe.

terminal presumably contributing to their effector function (Stam et al., 2013, Amaro et al., 2017). One important role of all effectors, CRN, RXLR, and others, is to suppress

the plant immune system, this is termed effector-triggered suppression (ETS) and is required for disease on host plants.

Plants have a second layer of immunity however, based on the recognition of effector genes (sometimes termed avirulence genes - AVR) from potential pathogens. This second recognition mechanism of the plant immune system additional to PTI is mainly intracellular in its action, and is based on either the direct, or more often indirect recognition of effectors, and is termed effector-triggered immunity (ETI). This recognition is primarily mediated by a class of receptors that contain a conserved nucleotide-binding domain (NB) and a leucine-rich repeat domain (LRR), often termed NB-LRR. Several NB-LRR proteins that confer resistance to *Phytophthora* have been characterised (Vega-Arreguin et al., 2017). These NB-LRR make up the large majority of what are termed resistance genes (R-genes), each with a matching effector AVR gene, which upon recognition triggers the immune response.

Continued development and layering of the immune response (ETI) and pathogenicity (ETS), this back and forth between host and pathogen is an evolutionarily given. Pathogen evasion of ETI is necessary for a pathogen to successfully infect its host plant. This can be achieved in various manners. It is known that pathogen effectors have high levels of allelic diversity, or that the pathogens have high levels of gene copies with high diversity to maintain an ever-evolving effector repertoire (Goss et al., 2013). In addition, effector genes often have a relationship with genetic elements that allow duplication and mutation. Having a diverse family of genes that will be under almost constant evolutionary pressure mean that pathogens often can lose genes that confer plant resistance (AVR genes) while maintaining pathogenic fitness, potentially even maintaining the pathogens ability to target the same host target. Avoidance of the plant immune system, i.e. ETI suppression, can also be achieved directly by

effectors (Staal et al., 2006). Consequently, plants need an ever evolving suite of R-genes to combat the innovation that takes place in microbes as evidenced by the rapid expansion and evolution of effector repertoires. This evolutionary back and forth between pathogen and plant, and the continuing arms race of effector and R-gene is described by the zig-zag model of plant immunity. The zig-zag model postulates that intense selective pressures drive the emergence and death of molecular innovations responsible for enhanced immunity and virulence in host and pathogen respectively (Figure 1.1).

1.2.2 The plant immune response

The dual system of PTI and ETI make up the plant immune system, and result in the activation of a suite of complex cellular, chemical, and molecular mechanisms that aim to limit infection and colonisation. The signalling pathways that are triggered as a result of PTI and ETI are well described (Dodds and Rathjen, 2010), utilizing kinase signalling, calcium signalling, NB-LRRs to trigger changing in transcriptomic, cell morphology and make up, and changes in the well-described salicylic acid and jasmonic acid–ethylene hormone pathways (Jones and Dangl, 2006). A key component of this system is the biosynthesis of small molecular secondary metabolites with antimicrobial activity that limit pathogen growth. Large numbers of diverse metabolites with a putative function in the plant immune system have been identified (Piasecka et al., 2015). For example, Glucosinolates (GS) are a group of secondary metabolites common to the Brassicaceae family, and Camalexin is a characteristic secondary metabolite found in *Arabidopsis thaliana*. These two in combination produce a disease resistance phenotype in *Arabidopsis thaliana* to *Phytophthora brassicae* (Schlaeppi and Mauch, 2010). Defensive strategies also ubiquitously include the accumulation of reactive oxygen species (ROS), these ROS

have multiple functions including signalling pathogen recognition, inducing cell death response, or having antimicrobial activity. The pathways that produce ROS are frequently the target for pathogen effector genes and therefore must be a key component of plant immunity (Jwa and Hwang, 2017). Often, however, the immune response results in localized cell death at the site of infection, known as the hypersensitive response (HR). The HR limits the progression of infection by limiting nutrient supply to the pathogen. All these together make an effective barrier to most pathogens, whether induced by PTI or ETI. To be pathogenic therefore microbes need to suppress this innate immunity.

1.3 Determinants of host range and host specific pathogenicity

Pathogenicity is the capacity of a microbe to cause disease. Host specific pathogenicity, therefore, is the capacity of a microbe to cause disease in a specific host species. What factors determine host specific pathogenicity and how these factors contribute to, and define the host range of a pathogen is still an open question. There will undoubtedly be many individual factors and mechanisms which enable pathogenicity on multiple hosts (i.e. enable a broad host range). Due to the variety of potential hosts and the fact that a single pathogen may be able to cause disease on a multitude of these distinct organisms it is safe to presume that factors that determine pathogenesis on a single host may not necessarily aid pathogenesis on another. Therefore a pathogen able to cause disease on a broad range of hosts must have a suite of host specific pathogenicity determinants to employ for each of those hosts. It is perhaps also the case that these individual determinates of host specific pathogenicity, taken together, do not encompass the entire mechanics that enable pathogenicity of multiple hosts. There is potential for other, general molecular

mechanisms that enable the maintenance of broad host range separate from multiple individual host specific pathogenicity determinants. What individual general mechanisms or sum of factors enable broad host range is again, very much, an open question. The themes and finding of the current research into the factors that determine host specific pathogenicity and enable broad host range have been explored here.

1.3.1 Host Range and Host jumps

Host range is generally defined as the subset of plant species that a certain pathogen species can cause disease on. There is an argument to be made, however, that this definition should be widened to include those plants on which a pathogen can proliferate regardless of symptoms. It is also important to distinguish between those hosts that the pathogen has been observed to proliferate on and cause disease in a natural environment, and those that have only be observed as hosts in a laboratory environment. In addition, there are what are termed non-hosts, plants that the pathogen is unable to colonise. Factors that determine a non-host/hosts status could include a successful PTI/ETI immune response, the presence or absence of essential nutrients for the pathogen, or simple that non-host is not present in the pathogens environment.

Regardless of the breadth or narrowness of the definition used, pathogens can be split up into two broad categories, specialist and generalists. The factors that determine the fitness of specialist (those able to cause disease in hosts from one or a few closely related taxa) versus generalist (those able to cause disease in hosts from multiple unrelated taxa) is complex and not well understood, moreover it is more a question of epidemiology and evolution and not molecular biology. Briefly however, a generalist

pathogen would gain greater opportunities for proliferation and spread, it is usually assumed however that evolution favours a specialist approach. There is largely assumed to be a fitness cost with breaking resistance in a new host and maintaining virulence on that host and the original (Bahri et al., 2009, Morris and Moury, 2019). Furthermore, different hosts embody different selective pressures, maintaining fitness across many hosts results in trade-offs limiting fitness to a single host (Kirchner and Roy, 2000). This pressure results in the specialisation of pathogens to a specific host. A clear example of this specialisation can be found in *P. infestans* isolates taken from France and Morocco, isolate from France were more aggressive on cultivars of potato common to France and less aggressive to those cultivars found in Morocco, the Moroccan isolates showed similar specialisation favouring Moroccan potato cultivars over French (Andrivon et al., 2007). However, generalist pathogens continue to thrive, apparently showing a reduction in cost for maintaining multiple host fitness or increased benefits from increased opportunities for proliferation and spread. It is important to investigate and understand the determinants of host range to understand how pathogens maintain multiple host fitness. These determinates could be generalised mechanisms that enable broad host range, or the cumulative effect of multiple individual host specific pathogenicity determinants.

Many studies that examine the determinants of host range and specific host pathogenicity concentrate on host jumps. A host jump starts when a pathogen encounters and infect a new host. Oftentimes, this interaction is not very productive for the pathogen and can result in extreme selective pressures acting on key (a)virulence factors. This results in adaptation of the pathogen to the new host, and is often followed by specialisation; a potential mechanism for speciation. However, it is necessarily true that some pathogen species are able to expand their host range

without experiencing the diversification that precedes and defines speciation. A diverse range of mechanisms have been attributed to functionality in host jumps, including hybridization, gene deletion, amino acid substitutions, and horizontal gene transfer (Morris and Moury, 2019). Hybridization between subspecies can be the impetus for a host jump, or maybe more specifically a host range expansion, for example, *Blumeria graminis f. sp. triticale,* was formed from the hybridization of a *B. graminis* subspecies specialised to wheat and another subspecies specialised to rye, resulting in a species which hosts include both wheat and rye (Menardo et al., 2016). Hybridization can also occur through distinct species hybridization, many *Phytophthora* species are known to have their origins in inter-species hybridisation, enabling host jumps and expansion of host range. (Depotter et al., 2016). One *Phytophthora* species with multiple *solanaceae* hosts known to have emerged through the hybridisation of *P. infestans* and an unknown species is *Phytophthora andina,* (Goss et al., 2011). Gene deletion has also been shown to be another mechanism that contributes to host jumps, extensive loss of genes, particularly putative secreted proteins, in the smut fungus *Melanopsichium pennsylvanicum* which is thought to have enabled jump from monocot hosts to dicot hosts (Sharma et al., 2014). In contrast, a single gene loss can also be the factor that allows a host jump, the loss of function of an AVR gene that was a target for specific R genes found in many wheat cultivars, allowed evasion of the wheat's immune system. And, is thought to be the onus for Wheat blast fungus's host jump to common wheat (Inoue et al., 2017). It is not just loss of genes that lead to host jump but also mutation in the form of amino acid substitution, for example *Phytophthora* species were able to jump between *Solanum* species via an amino acid substitution in a protease inhibitor (Dong et al., 2014).

Host jumps and host expansion can also occur when a new host species is introduced to the environment of the pathogen, or the pathogen is transported to a new area where potential hosts are located (Panstruga and Moscou, 2020). In such a situation the determinants of virulence for a single host may contribute to virulence on other potential hosts. What role effectors and other determinants of virulence may have in potential new hosts and non-hosts is a relatively understudied area of research (Stam et al., 2014). Some effectors have shown to have activity in non-hosts, for example, 33 *P. infestans* RXLR were tested for their ability to suppress FRK1 (FLAG22-INDUCED RECEPTOR-LIKE KINASE 1) in the *P. infestans* host tomato and the non-host Arabidopsis. 8 RXLRs were able to suppress FRK1 in tomato, whilst only 3 maintained this activity in Arabidopsis. Suggesting the 5 other effectors lost their ability to affect their target in the non-host plant (Zheng et al., 2014). Although in this example 3 effectors maintained effectivity in a non-host species, other studies show how key hosts determinants from closely related oomycete species were unable to suppress immunity in the other species' hosts and vice versa. *P. infestans* and *P. mirabilis* are close phylogenetic relatives, and have both evolved to specifically inhibit proteases from their distantly related respective hosts tomato and four o'clock flower (Mirabilis jalapa) (Dong et al., 2014). The *P. infestans* is a cysteine protease inhibitor PiEPIC and the homologue from *P. mirabilis* (PmEPIC1) whilst both shown to be active against the cysteine protease from the pathogens respective host, where shown to be inactive against the homologue from the other pathogens respective host. PiEPIC1 was able to suppress the tomato proteases RCR3 but not the Mirabilis homologue MRP2. And PmEPIC1 whilst able to suppress MRP2 had no activity against RCR3. It has been hypothesised that pathogens may be unsuccessful in the colonisation of distantly related non-host through the failure of effectors, specifically evolved for activity on the

host plants, to affect their targets (Dong et al., 2014, Antonovics et al., 2013). In fact, the phylogenetic relationship between hosts seems to be a key determinant in predicting host breadth and the probability of host jumps (Gilbert et al., 2012, de Vienne et al., 2013, Gilbert and Webb, 2007). Although there is little study as to why this may be the case, an obvious explanation would be the plants from similar plant families offer similar environments and resources. It could also be the case that related plants share aspects of the immune defence strategy whether that be specific R genes that have been circumvented, or more basic non-host resistance strategies e.g. recognition of PRR that the pathogen has lost (Schulze-Lefert and Panstruga, 2011).

1.3.2 Genomic Structure and Effector Repertoire Expansion as Determinants of Host Range

It is clear from our understanding of plant immunity, and the evolution of host range and the factors that contribute to host jumps, that the effector repertoire of a pathogen is key to determining that pathogens host range. A good example of this principle can be found in the filamentous fungi of the genus *Fusarium*, many of which cause blight, wilting and root rot in crops, but have marked differences in their host range. Three notable members are *Fusarium graminearum* and *Fusarium verticillioides*, important pathogens of cereal crops, and *Fusarium oxysporum* f. sp. *lycopersici* (*FOL*) which also infects monocotyledonous plant but can also infect dicotyledonous plants, is a notable pathogen of tomatoes, and boasts a much broader host range. Genomic analysis of these three fungi showed that *FOL* had four unique chromosomes representing over 25% of its genome. It was discovered that genes know to be key to tomato pathogenicity, Six1 (Avr3) and Six3 (Avr2), as well as an oxidoreductase (ORX1), known to be secreted *in planta*, were found on these chromosomes. More notable is the fact that partial or total transfer of these 4 chromosomes into a non-

pathogenic strain of *Fusarium oxysporum,* enabled that strain to infect tomato plants. Demonstrating that this distinct suite of genes was necessary for specific host pathogenicity (Ma et al., 2010).

Other genomic studies of filamentous pathogens reveal a similar process demonstrating the importance, in terms of virulence and pathogenicity, of a genomically distinct and diverse effector gene repertoire, and how this has evolved in those pathogens. Genomic analysis of six closely related *Phytophthora* species showed that gene-sparse areas of the genomes, those areas that are known to contain genes induced *in planta,* and especially effector genes, undergo rapid evolution. This is in contrast to gene dense areas which mostly contain genes common to all *Phytophthora* species and undergo less rapid evolution. Other families of genes found in these highly dynamic, gene sparse regions include several enzymes e.g. cell wall hydrolases and proteins involved in epigenetic maintenance. Another notable family was enriched in gene-spares regions, perhaps surprisingly was histone and ribosomal RNA (rRNA) methyltransferases, often showing presence/absence polymorphisms (Raffaele et al., 2010). This phenomenon has been observed in other filamentous plant pathogens and has been termed “the two-speed genome.” The pathogens genome is structured in such a way that pathogenicity factors undergo constant adaptive evolution in certain genomic areas, regions that aid gene expansion and diversification (Dong et al., 2015). Although this could be a mechanism by which pathogens maintain, and possibly adapt, the repertoire of pathogenicity factors that determine their host range, so observance of this so-called two-speed genome is not unique to broad host ranged pathogens. A similar genomic structuring also plays a key role in the host range of the fungal wheat pathogen *Zymoseptoria tritici*. Studies identified clusters of co-regulated genes, induced specifically during the infection of wheat compared to

another host. These pathogenicity islands showed signatures of positive selection during the evolution and specialisation of *Zymoseptoria tritici* and may represent determinants of host specific pathogenicity (Kellner et al., 2014). This study introduces a possible mechanism of dynamic adaption to the host, that is, co-regulation of pathogenicity factors specific to that host. Interestingly of those co-regulated genes, detoxification genes, particularly gene involved in peroxidase activity, oxidoreductase activity, and antioxidant activity were predominant and were key in determining pathogenicity on a specific host (Kellner et al., 2014). Representing a host-specific adaption and response to toxic secondary metabolites and reactive oxygen species produced by host plant immune system. This is evidenced by several cases of detoxification factors, like these, being factors that contribute to host specific pathogenicity in other fungal pathogens (Bowyer et al., 1995, Coleman et al., 2009, Srivastava et al., 2013).

The genomic structure of pathogens often includes portions that are diverse from other closely related species and a key to determining host specific pathogenicity, they often show hallmarks of evolutionary selection as well as the ability to create diversity and mutations. (Hacquard et al., 2013, Raffaele et al., 2010). It is also shown that these regions of diversity of pathogenicity islands are often differentially expressed as a form of host-specific dynamic adaption (Kellner et al., 2014). It is clear that the suites of effectors and other factors of host specific pathogenicity that a pathogen possesses define its host range. It is also clear that pathogens have evolved ways to expand and regulated the pathogenicity factor repertoires (Liang and Rollins, 2018).

1.3.3 Differential Gene Expression as a Driver of Dynamic Adaption to distinct Host Plants

A key example of this differential gene expression during distinct host infection can be found in the fungal cereal crop pathogen *Fusarium graminearum.* (Harris et al., 2016). Interestingly transcriptomic analysis revealed that during the infection of barley and wheat, the gene expression of *Fusarium graminearum* was similar in both host plants. However, when infecting maize, a third host, a group of differentially expressed genes was revealed. Showing the ability of *Fusarium graminearum* to dynamically adapt to different host may be due to the expression of a different set of genes. The importance of differential expression for host specific pathogenicity is also demonstrated within the *Phytophthora* species. Genomic analysis of *Phytophthora cactorum* an extremely broad host ranged pathogen of rhododendron and many other woody species, revealed expansion of detoxification enzymes, suggesting these are a key determinant of its broad host range. More specifically ATP-binding cassette (ABC) transporter families and major facilitator superfamily (MFS), as well as the cytochrome P450 (CYPs), peroxidase (POD), glutathione S-transferases (GSTs), methyltransferase (MTR), and dehydrogenase were all expanded in the *P. cactorum* genome. (Yang et al., 2018). Differential expression of these families of genes is part of several *Phytophthora* species response to the ginsenosides, notably *P. cactorum*, *P. capsici* differentially express 267 and 408 respectively when cultured with ginsenosides. (Yang et al., 2018). *Sclerotinia sclerotiorum* is another broad host range pathogen, studies of differential expression of genes in response to two host plants from different host families revealed host specific differential expression. However, in this species, the effect of distinct host plants on the transcriptome of the pathogen was limited, 53 of 628 genes. Although these genes did all have roles in the detoxification of host metabolites (Allan et al., 2019). Showing that whilst genomic structure and differential

expression may not be a universal determinant for host specific pathogenicity or dynamic adaption to multiple hosts, these concepts do seem to be a recurring theme in understanding broad host range. It is therefore fair to hypothesise that diversification and evolutionary selection of effectors genes and the subsequent expression of a subsets of the effector repertoire determines host range and dynamic adaption of a pathogen to a distinct host.

A crucial stage of a pathogens life cycle and surely a key component of a pathogens ability to infect multiple hosts involves its ability to perceive a distinct host and act accordingly. In some fungal pathosystems, G-protein coupled receptors (GPCR) play a key role in sensing environment and initiating downstream signalling pathways. One such GPCR, called PTH11, is found abundantly in the species of the fungal pathogen class *Sordariomycetes*. In the fungal insect pathogen *Metarhizium acridum*, multiple PTH11-related receptor paralogs were used to distinguish host from nonhost, however, in a closely related but more generalist pathogen, a single *Metarhizium roberstii* utilized a single PTH11-related receptor to recognize multiple hosts (Quandt et al., 2016). The GPCR class of receptors is not found in great numbers *Phytophthora* species, especially *P. capsici*. However, a mechanism of host perception must remain, this could be direct recognition through receptors similar to GPCRs in function, or indirect mechanism, such as nutrient sensing mechanisms like the nitrogen metabolite repression regulator (NMRA) which alter transcription as part of a potential nitrogen metabolite sensing mechanism (Pham et al., 2018).

1.3.4 Other Mechanisms of Host Range: Necrotrophs, and bacteria.

Pathogens can crudely be divided into two different groups, based on their strategy of infection and colonisation, those pathogens where the death of the host's cells happens in consort with, or even prior to, the colonisation of the host, these are called necrotrophs. In contrast, there are those pathogens that colonise the plants, whilst suppressing the plant immune response, maintaining host cell viability, and cell death occurs post colonisation, these pathogens are called hemibiotrophs. To some extend this division may be a false dichotomy, and even the notion that a neat line can be drawn to divide the two extremes of these infection strategies, biotrophic or necrotrophic, may be somewhat baseless (Oliver and Ipcho, 2004, Oliver and Solomon, 2010b). Even still, how different the methods and mechanisms of maintaining broad host range can differ in these two types of pathogens remains an open question. The majority of what is discussed above in this thesis may largely pertain to hemibiotrophs, however much many of the mechanisms for the maintenance of a generalised lifestyle and broad host rages may be shared amongst so-called hemibiotrophs and necrotrophs.

Some necrotrophic pathogens have remarkable broad host ranges *Botrytis cinerea*, for example, is a necrotrophic fungus with hosts from 146 taxonomic families, *Rhizoctonia solani* a soil-borne fungal pathogen has hosts in 169 families (Newman and Derbyshire, 2020). Additionally, distinct mechanisms, or determinants of host range have been proposed for necrotrophic pathogens. It has been noted that many taxa of broad host ranged plant pathogens display a greater amount of asexual reproductive habits than their narrow host counterparts (Gibson, 2019, Ross et al.,

2013, Newman and Derbyshire, 2020). The reasons for this seems unclear, some have theorised that Large populations enabled by a generalist lifestyle enables random mutations to allow for enough genetic diversity, and minimise the amount of deleterious mutations (Newman and Derbyshire, 2020). It has also been noted that in the insect fungal pathogen genus *Metarhizium*, that there is a significantly close relationship between the genome size and gene coding capacity of the species and the potential host range of the species (Hu et al., 2014).

In addition to an asexual lifestyle and expanded genomes, several specific molecular mechanisms, and strategies of virulence have been noted to be common in broad host ranged necrotrophic pathogens. These include the modulation and adaptive adjustment of host ROS and host pH (Newman and Derbyshire, 2020). In addition, it has been noted, as in hemibiotrophic pathogen, detoxification of a variety host-derived antifungal secondary metabolites is necessary for colonisation of a broad range of host plants. However as with Hemibitrophic pathogens, the evolutionary benefits of host generalism is not well understood, and it has been noted that more research into the molecular mechanisms that enable broad host range is necessary.

The discussions of host range above has largely been focused on eukaryotic pathogens, oomycetes and fungi. However, a large portion of the most devastating broad host ranged plant pathogens are bacteria. However broad theme in the mechanism of broad host range can be recognised. Bacteria genomes or often organised into sections of core conserved sequences and accessory genomic islands. These genomic islands encode a variety of organism specialisations, such as pathogenesis. A gene island can carry genes from diverse origins, built piece by piece through deletion and insertion events, they can also be exchanged between organisms

by lateral gene transfer. In pathogens, these genomic islands have frequently been found to encode virulence factors such as type III secreted effectors, these specialised genomic islands can be referred to as pathogenicity islands (Hacker et al., 1997), and characterisation of the genomic island has revealed host specific virulence factors ((Parkhill et al., 2001a, Parkhill et al., 2001b, He et al., 2004). Experiments have shown that virulence genes can be acquired by horizontal gene transfer of pathogenicity islands (Lindeberg et al., 2008, Lovell et al., 2009). Diversification of effector repertoire has also been noted in bacterial strains through homologous recombination, contributing to differences in the host range of different strains (Yan et al., 2008), in fact in large part it seems the gene – for gene hypothesis of plants immunity and by extensions the acquisition and loss of key virulence genes such as genes coding for type III secreted effector proteins and key in defining and expanding host range. These genes either working as important avirulence factors, or key effectors determining virulence (Lindeberg et al., 2008, Cai et al., 2011). The presence and absence of certain PAMP triggers has also been shown to be a key determinant of host range. For example, different flagellin alleles from strains with different host ranges, and host range breadth, has a marked effector on the bacterial growth in planta (Takeuchi et al., 2003). However, it is still up for debate how important the loss or gain of certain key effectors, or PAMP triggers is in determining host range evolution (Cai et al., 2011).

1.4 A highly destructive broad host ranged plant pathogen: *P. capsici*

Phytophthora capsici is a highly virulent and versatile pathogen of crop plants that occurs throughout warmer climates (Lamour et al., 2012b). *P. capsici* was isolated from diseased chilli peppers in New Mexico and described in detail for the first time in

1922 (Leonian, 1922). It is a filamentous oomycete plant pathogen that causes crop blight, and can infect plants at any stage of development, causing seedling damping-off, crown root and foliage blight, and fruit rot prior or post crop harvest. Its host range encompasses multiple members of the *Cucurbitaceae* or gourd family, and the *Solanaceae* or nightshade family, major crop hosts include but are not limited to peppers (*Capsicum annuum*), cucumbers (*Cucumis sativus*) squashes (*Cucurbita maxima, Cucurbita moschata, Cucurbita pepo*), gourds (*C. moschata*), watermelons (*Citrullus lanatus*), melons (*Cucumis melo*), tomatos (*Lycopersicon esculentum*), eggplants (*Solanum melongena*), and black pepper (*Piper nigrum*) (Figure 1.2). More recently identified hosts include green beans and lima bean (Babadoost et al., 2008), in total studies suggest that its host range could include up to 49 different species of crops and weeds (Drenth and Guest, 2004, Erwin and Ribeiro, 1996, Tian and Babadoost, 2004).

In the United States of America in 2019, approximately USD$621 million worth of chilli and bell pepper were grown, using 48,500 acres of land (NASS, 2020) all at threat from *P. capsici* disease. Taking into account the broad range of crop plants that *P. capsici* infects, global food crops under threat each year are estimated to be worth over USD$1 billion (Lamour et al., 2012b). In addition to the economic effect, pathogens such as *P. capsici* can have a humanitarian impact. For example, in Sarawak state, Malaysia, where over 60,000 family farms grow black pepper, making it the most important crop in that region, a *P. capsici* outbreak in the mid-1950s caused 100% crop losses throughout the whole state (Drenth and Guest, 2004), and the state is still blighted by the pathogen today with one region suffering 75% disease incidence (Farhana et al., 2013). Necessarily having ramification for the ability of the farmers to provide for themselves and their families.

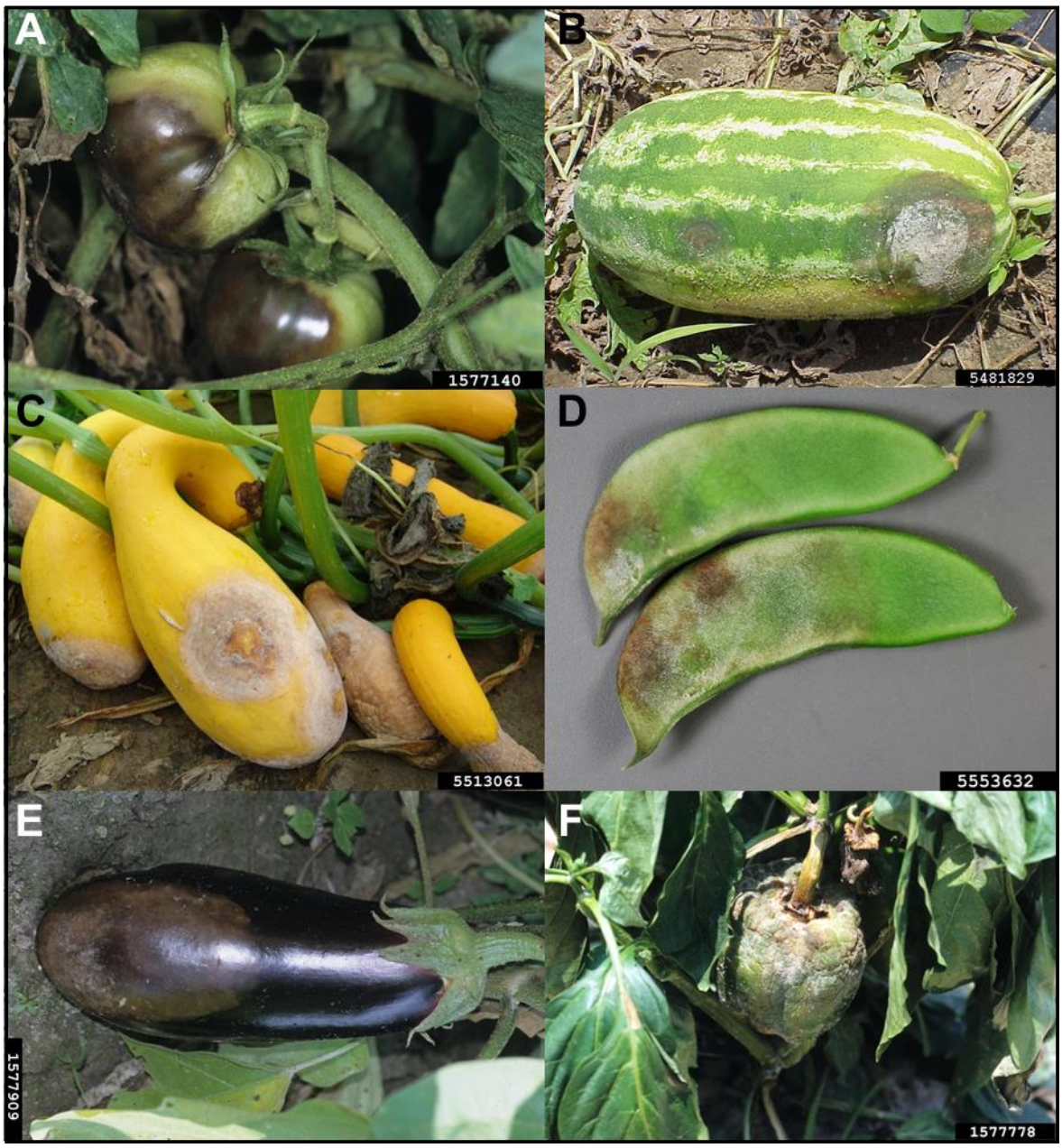


Figure 1.2: Infection in the field of multiple host plants.

(A) *P. capsici* infection of a garden tomato cv. (*Solanum lycopersicum*) displaying buckeye rot of fruit. (B) *P. capsici* infection of watermelon (*Citrullus lanatus*) showing large lesion and sporulation. (C) *P. capsici* infection of on fruit of yellow crookneck squash (*Cucurbita pepo*) showing large lesion, sporulation and fruit root. (D) Signs of *P. capsici* infection on lima bean (*Phaseolus lunatus*) in a laboratory setting. (E) *P. capsici* infection of eggplant (*Solanum melongena*) displaying a large lesion and the beings of sporulation at the base. (F) *P. capsici* infection of pepper showing extensive fruit rot and sporulation (*Capsicum annuum*). Pictures A, C, E and F courtesy of Gerald Holmes, Strawberry Center, Cal Poly San Luis Obispo, Bugwood.org, picture B courtesy of Jason Brock, University of Georgia, Bugwood.org, picture F courtesy of Nancy Gregory, University of Delaware, Bugwood.org

The broad host range and global reach, while not unique amongst *Phytophthora* species, gives *P. capsici* a far-reaching and profound impact on crop growers. However, the studies of, and molecular tool for the studies of, other *Phytophthora* species such as *Phytophthora infestans*, or *P. sojae*, both of which have relatively small host ranges, are further advanced. Despite the ease and speed it can be cultured in laboratory setting *P. capsici* remains understudied. Whilst a partial genome was published in 2012 (Lamour et al., 2012a) genomic tractability remains elusive. However, *P. capsici* is well-placed to become a model organism for studying the determinants of host specific pathogenicity and range, and the mechanisms of dynamic adaption to host in oomycetes and other pathogens. Therefore, this thesis aims to expand our understanding of the mechanisms that underlie broad host range in *P. capsici* and develop tools to aids in the study of oomycete plant pathogens, specifically increasing genetic tractability and the ability to conduct transcriptomic studies during key but recalcitrant stages of a pathogens lifecycle.

1.4.1 *P. capsici* as an evolutionary power-house

P. capsici can reproduce both sexually and asexually, and genetic variation, in particular deviations in host range, are well reported (Tian and Babadoost, 2004). *P. capsici* is a heterothallic species and, therefore, in the field two mating types (A1 and A2) exist and are required for sexual reproduction. Mating leads to the formation of thick-walled oospores that can remain dormant in the soil for long periods and persist through harsh environmental conditions (Figure 1.3). Because of their persistence and longevity, the presence of oospores in the soil can render fields unsuitable for the production of susceptible crops. Sexual reproduction also means that *P. capsici*

populations are genetically diverse. These features, alongside high mutation rates in *P. capsici* when subject to UV irradiation, can lead to fast shifts in population structure and fungicide resistance (Barchenger et al., 2018). In fact, *P. capsici* is notoriously malleable and demonstrates rapid adaptability, displaying a single nucleotide variation (SNV) density and diversity significantly higher than that of most eukaryotes. Further genomic malleability is demonstrated by the remarkable high instances of loss of heterozygosity, where tracts of the genome from one of the parents is switch to the other parent's haplotype, allowing mating-type switching (Lamour et al., 2012a). It is therefore widely assumed that genetic diversity underpins the remarkable adaptability of *P. capsici*.

Under favourable field conditions, *P. capsici* is also able to reproduce rapidly by asexual reproduction. Following successful infection, large numbers of sporangia are produced on the surface of plant stems, leaves and fruit. Over three billion viable spores have been reported on the surface of a single squash fruit (Lamour et al., 2012b). These characteristically lemon-shaped or limoniform sporangia (Figure 1.3) are induced to detach on contact with water, under humid or rainy conditions facilitating mechanical spread. They each contain, and thus release 20-40 free swimming, biflagellate zoospore, which acts as one of the main means of infection and spread in a field (Figure 1.3). These zoospores are chemotactically attracted toward plants and swim upwards being negatively geotropic. Remarkable they have been reported to survive in tail-water for up to 45 days (Roberts et al., 2005). Individual analysis of these zoospores has recently revealed a dramatic plasticity to the *P. capsici* genome; sequencing of single zoospores reveals extreme aneuploidy, termed Dynamic Extreme Aneuploidy (DEA), it means asexual reproduction results in high

intra-genomic variation in chromosome copy number varying from diploid (2N) to anywhere up to 6N or beyond (Hu et al., 2020).

Alongside meiosis and frequent outcrossing, loss of heterozygosity, high SNV and this dramatic genomic restructuring occurring during asexual reproduction, all happening in a potential three billion spores, each with up to 40 viable zoospores, the genetic diversity even in a single field of infected plants could be massive.

Once a zoospore, oospore or sporangia has found a host plant and germination has begun it has to invade the host plants tissues and successfully evade or suppress the host's innate immunes system. On contact with a plant, the zoospores encyst and germinate, these germinating cysts produce germ tubes that either directly penetrate the plant cuticle, which can take up to an hour (Hausbeck and Lamour, 2004), or can utilize natural opening such as stomata (Katsura and Miyazaki, 1960). Penetration is often aided by the enzymatic breakdown of the plant tissue (Yoshikawa et al., 1977). Direct germination of oospores and sporangia is also possible, and the disease progresses similarly.

1.4.2 The Infectious life cycle of *P. capsici*

P. capsici has a hemibiotrophic lifestyle. This is characterized by two distinct phases during the disease cycle. The first stage involves the development of an appressorium and penetration peg for invasion. Followed by hyphae, filamentous tubes, which colonise the plant tissue, and the formation of haustoria, structures that branch off from the hyphae and expand inside the plant cell walls creating an intimate shared membrane for the transfer of water, nutrients and effectors (Bozkurt and Kamoun, 2020)(Figure 1.3). Notably, this biotrophic stage includes the suppression of the plant immune response whilst the pathogen invades the host tissue. Specifically, in *P.*

capsici biotrophy is from 0 hours post inoculation (hpi) to about 16 hpi. It can tentatively be characterised as the portion of time when the haustoria formation marker gene *P. capsici* haustorium-specific membrane protein 1 (PcHmp1) is up-regulated and before the up-regulation of the necrotrophy marker *P. capsici* necrosis-inducing *Phytophthora* protein 1 (PcNpp1). This biotrophic phase (0 – 16 hpi) features the enriched induction of genes with roles in the expression and translation of genes, and protein metabolism. This demonstrates that in the early stages of *P. capsici* infection proteins are derived from *de novo* synthesis rather than from host derived amino acids (Jupe et al., 2013). This is consistent with the high expression of amino acid biosynthesis genes found in early *P. infestans* infection (Grenville-Briggs et al., 2005). And suggests a host perception induced bootstrapping of infection initiation and places a particular onus on the ability of germinating cysts to detect and adapt to the host and invade its tissue without deriving nutrient from the host.

Following this initial colonisation and growth inside the plant, the second phase is the necrotrophic phase, which involves the death of the plant tissues and continued nutrient harvest. The Necrosis and ethylene-inducing peptide 1 (Nep1)-like protein superfamily, a family of necrosis-inducing peptides key in pathogenicity, and genes associated with catabolism and degradation are found co-regulated with the necrotrophy marker gene PcNPP1. Induction of these genes are part of a transcriptional shift that occurs around the 16 hpi mark and represents the shift from biotrophy to necrotrophy. Following successful infection and during the death of the plant (or in a similar vegetative stage during the routine maintenance of lab strains) *P. capsici* grows as coenocytic filamentous mycelium, formed of multiple free branching hyphae, from which sporangia or oospores can be derived or induced (Figure 1.3). These life cycle phases while dynamic are closely regulated by shifts in

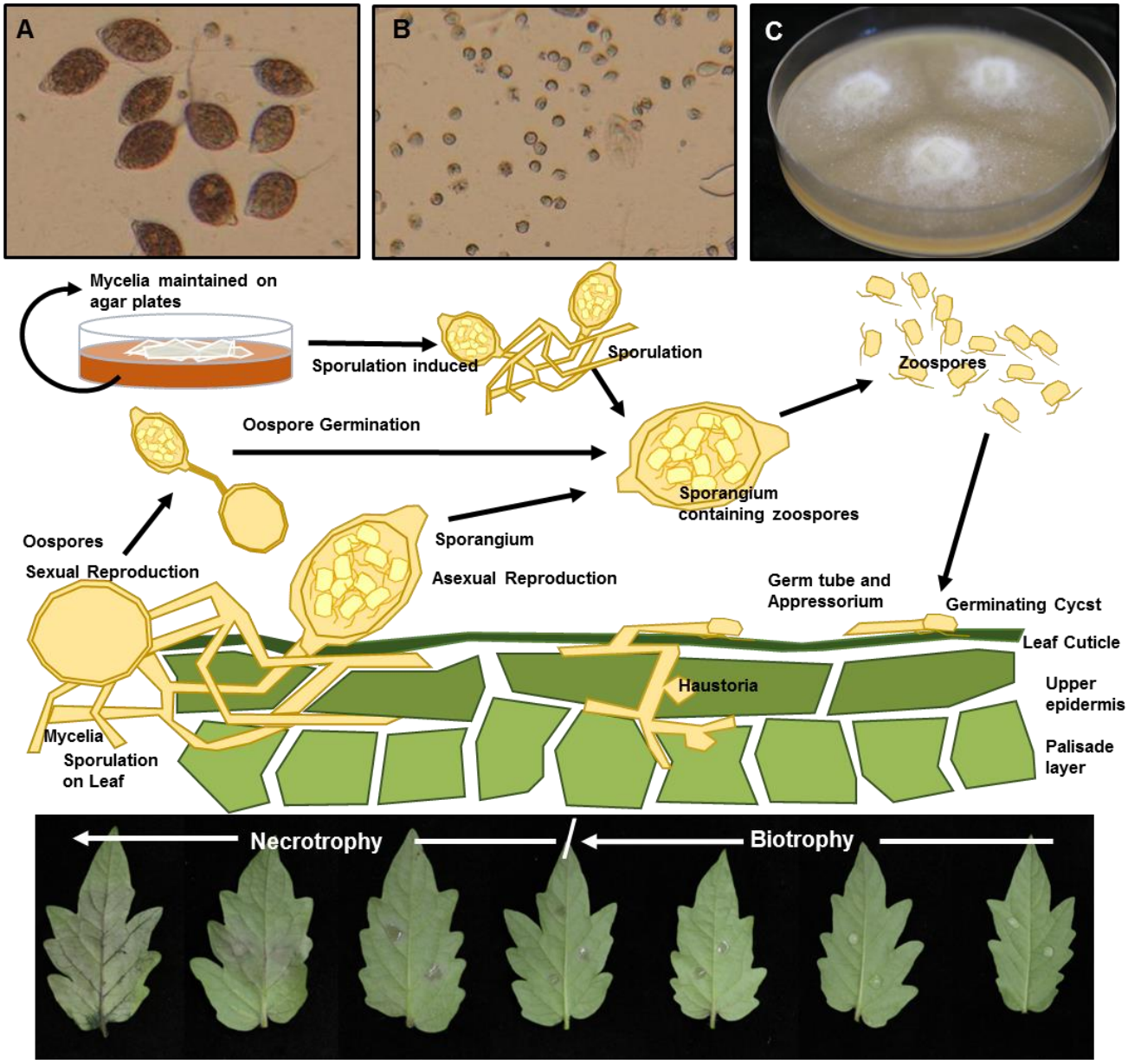


Figure 1.3: Life cycle images and diagram.

A A-C show images of P. capsici at various stages: sporangia filled with zoospores and empty (A), biflagellate swimming zoospores (B), and mycelia grown on V8 agar medium in the laboratory (C). The diagram shows the life cycle of P. capsici. Following colonisation and completion of the plant infection sporulation is induced and spores erupt from the plant tissue. Two kinds of spores are produced, oospores through sexual reproduction and sporangium through asexual reproduction. Oospores can persist for many years in the soil, but will eventually germinate and produce sporangium of their own. Sporangium production can also be induced in the laboratory. In the laboratory mycelia is maintained on V8 agar medium, and sporulation is induced through fresh plating and a specific dark-light cycle, which allows collection and isolation of sporangium. Sporangium contain up to 40 biflagellate mobile zoospores which are released. These are the main inoculum material for *P. capsici.* Upon contact with plant tissue the zoospore encysts and germinates, forming

a germ germinates, forming a germ tube and an appressorium penetration structure. In the early stages of colonisation the immune system is inhibited and the plant tissue is not damaged. Haustoria form pushing into the plant cell, without penetrating the membrane, allowing the pathogen to form an intimate connection with the host cells. This is known as the biotrophic stage of infection. After colonisation of the tissue P. capsici moves into the necrotrophic stage of its lifecycle, drawing nutrients from the plant and causing water soaked lesions, which eventually give rise to sporulation and continuation of the cycle. Pictures A, B courtesy of Elizabeth Bush, Virginia Polytechnic Institute and State University.

the pathogens transcriptome. During the final stages of infection and upon sporulation, there is an increase in signalling and developmental processes (Jupe et al., 2013).

A lifecycle that includes a portion that is biotrophic and a portion that is necrotrophic is described as hemibiotrophic. The hemibiotrophic life cycle is common to many fungal plant pathogens such as *Magnaporthe, Colletotrichum*, and *Mycosphaerella* genera (Perfect and Green, 2001). Traditionally a biotrophic relationship has been considered sophisticated and thus advanced in terms of evolution and molecular mechanisms (Lewis, 1973). Biotrophy is thereby, characterised by complex molecular mechanisms aimed at fine-tuning the plant immune system. It enables colonisation without an immune response whilst maintaining the plant's ability to grow and thrive. During the biotrophic stages *P. capsici* forms intimate and complex morphological structures, such as haustoria, that invade the cell and extracting nutrients whilst also inhibiting cell death and the immune response. Thus, biotrophy seems to be a highly sophisticated co-evolved relationship. However, it is now clear that highly evolved, and complex host-specific killing mechanism are employed by necrotrophic pathogens as well (Oliver and Solomon, 2010a). It, therefore, seems remarkable that pathogens can maintain both biotrophic and necrotrophic lifestyles in the same host with fine-tuned temporal control as seen in *P. capsici* and other hemibiotrophs. Hence, a broad host ranged hemibiotrophic pathogen with a co-evolved, intimate and presumable

idiosyncratic relationship with multiple diverse host plants must have sophisticate pathogenicity strategies indeed. Identifying and understanding these mechanisms is one of the main goals of this research.

1.4.3 Dynamic host adaption in P. capsici

Previous studies have shown that the *P. capsici* infection progression phenotype is similar in four different host plants (Unpublished Huitema Lab). The progression of infection was examined in 4 host plants, the model plant *Nicotiana benthamiana* and three crops, cucumber (*Cucumis sativa*), tomato (*Solanum lycopersicum*) and pepper (*Solanum capsicum*). Macroscopic observation of inoculated leaves shows that *P. capsici* appears to grow similarly on all four host-species (Figure 1.4). Initially infection displays asymptomatic biotrophic growth. After approximately 16 hpi, surface mycelium is visible followed by necrotic, water-soaked lesions, which are forming at 32 hpi, and clearly observed at 72 hpi. Confocal microscopy on all four hosts at 16 hpi showed haustoria in all hosts (Figure 1.4 - white arrows). As Biotrophic infection seems established in all four host by 16 hpi this time point was chosen for transcriptomic analysis. The purpose of which was to examine if, despite the similar disease phenotype, differential gene expression was driving host specific pathogenicity.

Microarray experiments followed by differential gene analysis revealed 2452 genes are differentially expressed between the four hosts ($p<0.05$, BH correction) (Unpublished Huitema Lab). With 1246 genes significantly expressed with $p<0.01$. Using the more stringent significance about 13% of genes could be host dependently differentially expressed by *P. capsici*. Pairwise comparisons showed that cucumber, the single *Cucurbitaceae* host, is most different from all other hosts. Cucumber compared with any of the other hosts gives 1529, 2043 and 1850 genes that are

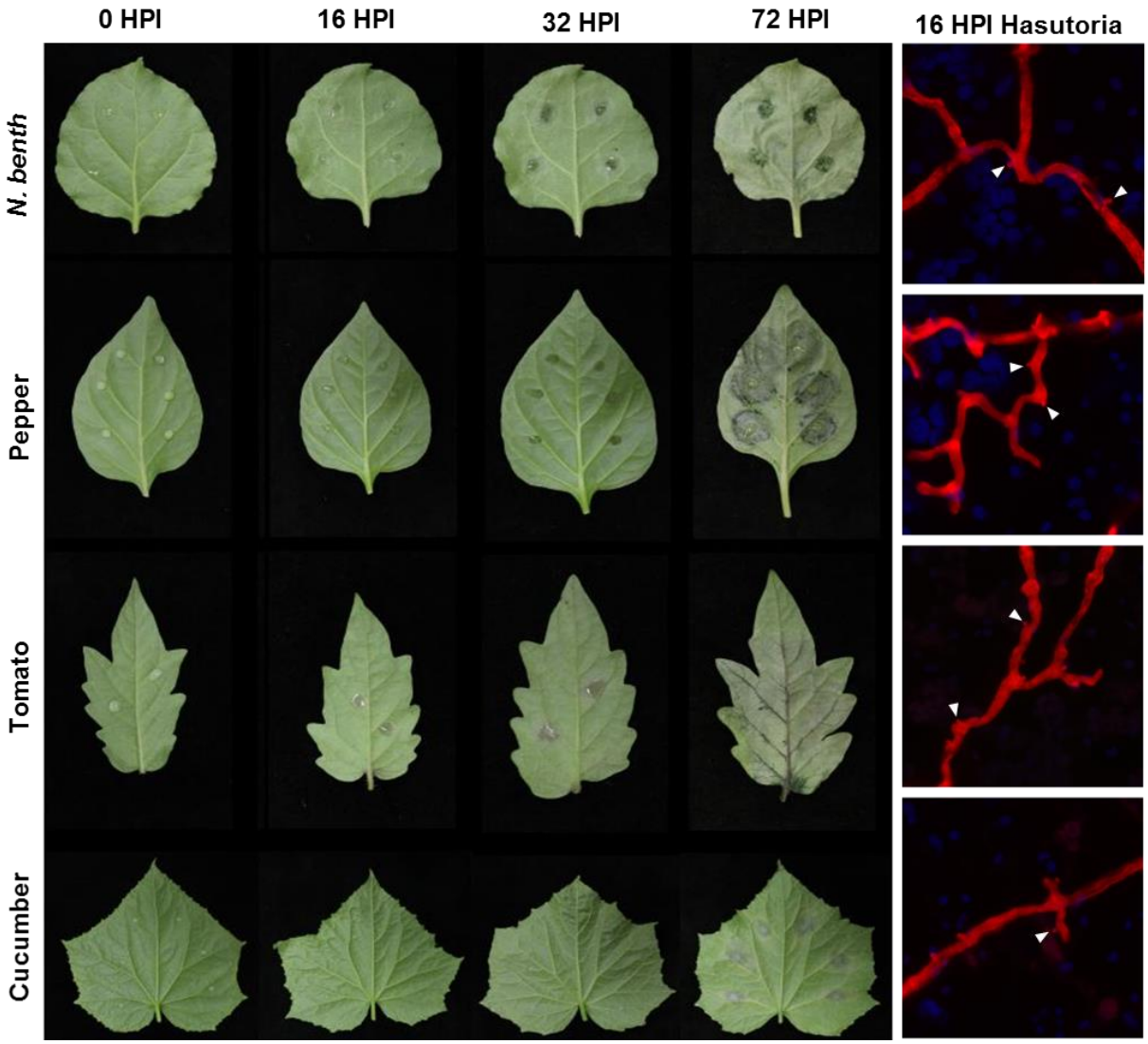
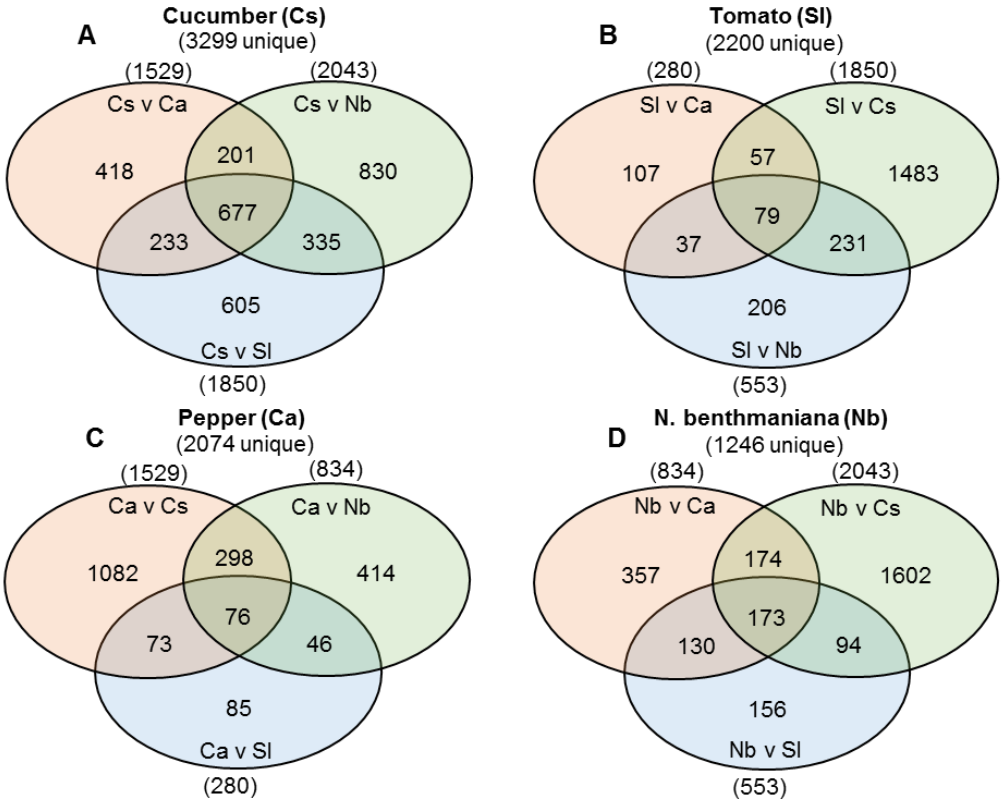


Figure 1.4: Four host leaf infection assay and confocal microscopy.

(Huitema lab unpublished results). Droplet inoculation of four *P. capsici* host plants shows the progression of disease. 0 – 72 hours post inoculation (hpi). 4 host plants have been used the model plant *Nicotiana benthamiana* and three crops, cucumber (*Cucumis sativa*), tomato (*Solanum lycopersicum*) and pepper (*Solanum capsicum*). Intial gowth 0 – 16 hpi in biotrophic and therefore relatively asymptomatic. However post 16 hpi symptoms in the form of mycelia can be observed in all for hosts, with water soaked lesion appearing in all four host by 72 hpi. Confocal microscopy at 16 hpi revels a similar infection phenotype across the hosts. Infiltrating hyphae and haustoria can be seen in red, with haustoria formation in all 4 hosts by 16 hpi (indicated with white arrows).

differentially expressed between cucumber and pepper, *N. bethamiana* or tomato respectively. Furthermore, when comparing pepper and tomato, the two most closely related solanaceae species, only 280 genes are differentially expressed.



E

	Cs v Nb	Cs v Ca	Cs v Sl	Nb v Sl	Ca v Sl	Ca v Nb
Total DE	**2043**	**1529**	**1850**	**553**	**280**	**834**
Up-Regulated	903	844	974	240	150	459
Down-Regulated	1140	685	876	313	130	375

Figure 1.5: Venn diagrams of the number of DE genes in each of the four host compared to each of the other hosts from Microarray experiments of *P. capsici* 16 HPI from leaf infection assay of four host plants.

The 4 host plants that have been used are the model plant *Nicotiana benthamiana* and the three crops, cucumber (*Cucumis sativa*), tomato (*Solanum lycopersicum*) and pepper (*Capsicum annuum*). (A) shows the 3299 DE genes of cucumber infection, (B) the 2200 DE genes of tomato infection, (C) the 2074 genes from Pepper infection, and (D) the 1246 DE gene of the *N. benthmaniana* infection all separated into the comparisons between each of the three other host plants. The table (E) shows each number of DE genes from each comparison group in its entirety with the total number of up regulated genes and the total number of down regulated genes.

Venn diagrams of the host overlap of differentially expressed genes were constructed to see if the same genes are differentially expressed in 2 or more host compared to a single other. Again, cucumber stands out as both prompting the largest number of uniquely differentially expressed gene (3299), and having the largest proportion of

those differentially expressed genes (677, 20.5%) found shared in the individual comparisons with each of the other 3 hosts (Figure 1.5 A). The other three host *N. benthamiana*, pepper, and tomato only having 173 (6.4%), 76 (3.7%), and 79 (3.6%) of their uniquely differentially expressed genes that are jointly different in a comparison with the other three hosts (Figure 1.5).

Within those genes that are differentially expressed RXLR effectors are amongst the most dramatically differentially expressed genes across all the host, in terms of the largest differences in transcript level (Unpublished Huitema Lab). In total 147 out of 463 (31%) RXLR appear to be regulated in a host-dependent manner, with a majority showing dramatic changes in gene expression. 26% of the top 200 most differentially expressed genes, in term of changes in transcript expression, are RXLRs. In the top 100, that number goes up to 31% and the top 10 most differentially expressed genes contain six RXLR genes: PcRXLR 373, 374, 030, 039, 475 and 325. This suggests that RXLR effectors can be switched on or off specifically dependent on the host being infected.

The principal expression profiles were identified by performing hierarchical clustering of all differentially expressed genes followed by the plotting of the genes expression profile (Unpublished Huitema Lab). These analyses resulted in the establishment of 10 distinct clusters, defined by host-dependent changes in expression. Within these clusters, 10 different expression profiles identified, ranged from host specific clusters to groups of genes that show expression on 3 out of four hosts (Figure 1.6). Cluster 1 and 4 show specific high expression on cucumber, whereas cluster 8 shows the opposite pattern. Cluster 3 shows higher expression in tomato and pepper and clusters 5, 6 and 7 are smaller clusters with higher expression in pepper, tomato or *N. benthamiana* respectively. The two largest are clusters 1, those genes that are

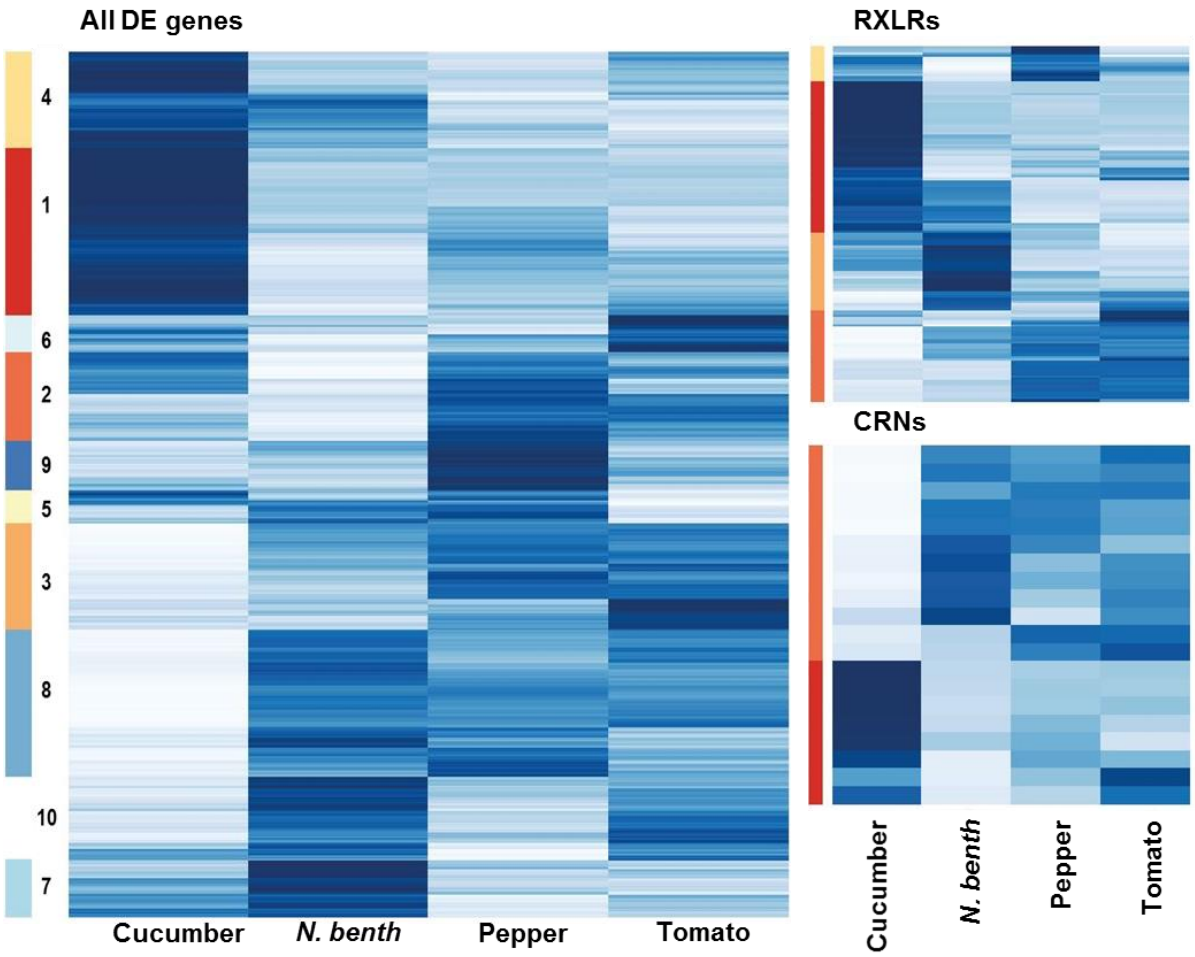


Figure 1.6: Heat map of Hierarchical cluster analysis of DE genes from Microarray experiments of *P. capsici* 16 HPI from leaf infection assay of four host plants.

The 4 host plants that have been used are the model plant *Nicotiana benthamiana* and the three crops, cucumber (*Cucumis sativa*), tomato (*Solanum lycopersicum*) and pepper (*Capsicum annuum*). Hierarchical cluster analysis shows 10 different expression profiles of all differentially expressed (DE) genes. Blue are those genes induce and white and those gene turned off. Hierarchical clustering of DE RXLR and CRN effector genes was also carried out. The RXLRs cluster into 4 expression profiles and the CRNs cluster into two.

induced in cucumber and down-regulated in the other 3 host plants, and cluster 8 whose genes are induced in the *Solanaceae* hosts and down-regulated in cucumber. When specifically clustering only the RXLR genes, four possible expression profiles 22 shows higher expression in both tomato and pepper and a third shows up-regulation in the more distal *solaneceous* species *N. bethamiana*. Cluster four is a very

small cluster that shows minor up-regulation in Cucumber and pepper, but not in the two other hosts.

These studies (Unpublished Huitema Lab) show that differential gene expression may be a driver of dynamic adaption to the specific host plant. It also supplied a list of potential genes that may be determinates of host specific pathogenicity and therefore host range, and that analysis of these genes may reveal the mechanism that allows maintenance of pathogenicity on a broad range of hosts. Analysis of the expression profile of these genes shows that of the four hosts tested, 3 being members of the *Solanaceae* family *Nicotiana benthamiana*, Pepper (*Capsicum annuum*), and Tomato (*Solanum lycopersicum*), the fourth host being cucumber (*Cucumis sativus*) of the family *Cucurbitaceae*, that the single *Cucurbitaceae* host plant induced the most differential expression of genes. This shows that differential gene expression may be a mechanism of dynamic host adaption of *P. capsici*, information that may illuminate the mechanisms of broad host range in oomycete. It also demonstrates that there may be an association between the genes being expressed and the closeness of the relationship of the host plant. That is, closely related hosts have more genes in common that are necessary for host specific pathogenicity, whereas, distantly related hosts have less of these shared host specific pathogenicity genes. Therefore, perhaps a better term for some of these genes may be determinates of genus specific pathogenicity rather than host specific pathogenicity.

1.5 Tools for the molecular investigation of pathogens

As previously explained much research into the mechanism of plant immunity are conducted in model species. These species are often genetically tractable and key insights into the fundamentals of pathogen-host systems can be gleaned from these

kinds of experiments, however, the applicability of these studies to inform resistance breeding or agricultural strategies to limit crop loss may be restricted. Novel techniques are constantly being developed across multiple biological disciplines which allow deep interrogation and smarter experimental design. However the ability to apply certain techniques to non-model organisms is limited, their implementation can be a great boon to picking complex processes and important biological problems. In particular developing tools in real-world relevant pathosystems, such as *P. capsici* may lead to a better understanding of broad host pathogenicity and dynamic adaption that is notable to that species and other key pathogens.

1.5.1 Gene editing and Characterisation

An obvious example of novel techniques with broad applicability that has been developed in recent time is the CRISPR-Cas9 nuclease gene editing tool. The methodology is based around the Cas9 nuclease that forms a complex with a sequence specific targeted short guide RNA. This complex can recognise and bind to the target genomic DNA via Watson-crick base pairing, the nuclease then selectively make a double stranded cut in the portion of targeted DNA. Endogenous DNA repair mechanism can introduce mistakes when joining the two DNA ends back together, termed non-homologous end joining (NHEJ), it can introduce base pair addition and deletion causing frameshifts knocking out the gene in a process called non-homologous recombination. The introduction of a repair template into this system allows specific editing, such as the insertion of specific and desired point mutations, or the introduction of large genes, like selections markers at the point of cleavage. The ability to knock out and edit, with a high level of accuracy, genes *in situ* has enabled a greater understanding of gene function in many biological systems. This system was first successfully utilized in mammalian cells in a pair of papers published in 2013 and

since has been in thousands of lab in various organism and system (Hsu et al., 2014). It was another 5 year until the first *Phytophthora* species *(P. sojea)*, was modified through genome editing (Fang and Tyler, 2016, Fang et al., 2017). This provided an important tool for the molecular investigation of oomycete plant pathogens. However, implementation in other oomycete species has been slow progress. A transformation procedure and functional editing of key genes have since been published in *P. capsici* (Wang et al., 2018, Miao et al., 2018) and functional studies have also been carried out in *Phytophthora palmivora* (Gumtow et al., 2018). However, several systematic attempts to implement the same system in *P. infestans* (van den Hoogen and Govers, 2018) has had no success, meaning the success of this system could be very species or even strain-specific.

1.5.2 “Omics” studies for functional analysis.

Over the last decade the advent of high throughput technologies, bioinformatics and increased computational power, and unbiased systems biology approaches has dramatically increased the ability to conducted broad “omics” based studies. This has had a vast impact on our understanding of complex processes in disease progression and lifecycle, host adaptation, virulence, and gene/protein function. In addition integration of multiple “omic” technologies can be a powerful technique for the discovery of system-scale mechanisms of pathogenesis and host adaption, and the molecules with key roles in pathogenesis (Ball et al., 2020). Proteomic and transcriptomic studies have been used to elucidate key genes and mechanisms for many lifecycle stages of *Phytophthora* species (Resjö et al., 2017, Pang et al., 2017). Or, in discovering key elements of fungicide resistance (Pang et al., 2016, Ma et al., 2018), and host adaptation in plant pathogens(Yang et al., 2018). However, due to a

large part of the lifecycle of the pathogen taking place inside the host plant, transcriptomic analysis of these key infectious stages remains elusive. The use of normal techniques to probe such time points mean that pathogen reads have to be distinguished from host reads which will always be in excess, this comes at the sacrifice of breadth and depth of sequencing, meaning the loss of low expression transcripts and difficulty in identifying all but the most differential of expression. This is a significant problem especially in the early stages of infection where pathogen biomass is so low.

Therefore, there have been many experimental designs to avoid the need to discriminate pathogen and host genes post-sequencing. Isolation of host genes can be conducted at any stage in the experimental design. There are two main types of experimental design that allow this, 1) *In vitro* systems and 2) Nucleic acid purification systems, these methodologies both allow separation of the pathogen material from that of the host, *in vitro* systems prior to RNA extraction and Nucleic acid purification post RNA extraction.

Examples of the former include, what are termed here as “sandwiches”, which are mycelia placed on the abaxial side of the leaf and a second leaf placed on top. This methodology has been used to characterizing those key genes of *P. sojae* of the early stages of soybean infection (Chen et al., 2007). A similar methodology has since been utilized to examine the interaction between *P. capsici* and the non-host model plant *Arabidopsis thaliana* (Ma et al., 2018). Other methods that have been trialled is culturing *P. palmivora* in different plant broth to detect novel secreted proteins (Apinya et al., 2018). Lastly, host interactions with the culturable tomatoes host cell MsK8 and several *Phytophthora* species have been investigated. Interestingly, only species pathogenic to tomato could infect the cultured cells and hallmarks of the plant immune

response were shown, reactive oxygen species and cell death response. Additionally, signatures of infection such as germ tube penetration and haustoria formation, as well as RXLR and other effector expression were observed (Schoina et al., 2017).

In contrast, there are methodologies that have been developed to involve purification of extracted RNA following normal *in planta* infection assays. Two popular methodologies that could fit this mould are translating Ribosome Affinity Purification (TRAP) Followed by RNA Sequencing Technology (TRAP-SEQ), and Pathogen enrichment sequencing (PenSeq). TRAP-SEQ involves the tagging of small conserved ribosomal proteins, purification of translating ribosomal complexes with their attached mRNAs using this tagged protein. Followed by isolation of the mRNA followed by sequencing. This has been used to isolate mRNA from rare cell types and tissue-specific cells in both mammalian and plant systems (Heiman et al., 2008, Reynoso et al., 2015), it could be utilized to isolate pathogen cells from infected plant tissue, however to date this has not been conducted. A different methodology that has been specifically designed to enrich pathogen DNA is PenSeq. PenSeq involves the creation of a library of genes that are to be enriched, incubating that library with the isolated DNA, binding of the tagged library to the complementary sequences thus allows purification and enrichment of these target sequences. Although, originally this methodology was developed for use with gDNA, it has since be used to investigate cDNA and thus has greater utility now for transcriptomic analysis (Lin et al., 2020).

These methodologies represent a wide range of options to use an “omics” style approach to interrogating the molecular mechanism of pathogenicity is key, but recalcitrant, and therefor understudied, infectious stages of the pathogen life cycle. Additionally, the implementation of the CRISPR-cas9 system in oomycetes is a significant step in increasing the genetic tractability of these species. It is thus, one of

the aims of this project to further explore and develop tools for broad omics style experiments, and advance powerful gene characterisation technologies in *P. capsici*.

1.6 Conclusion and Aims

P. capsici is a global pathogen with a far-reaching impact. Its oversized impact is perhaps largely due to its broad host range, meaning it can decimate cucumber and squash crop in the Midwest of America whilst impacting chilli pepper and black pepper farmers in Indonesia and Southeast Asia. Understanding what molecular determinates and processes enable its broad host range, therefore, is key to treating this disease on the global scale where it functions. Like other generalist pathogens, and like many oomycetes *P. capsici* has a remarkable penchant for genomic malleability and genetic diversity. The ability for genomic restructuring and expansion seems to be key in broadening and maintaining host range and is perhaps the impetus for host jumps in many species. Having a diverse or expansive genome is not the full story however, regulation in the form of differential expression of key genes seems to be a fundamental mechanism in the dynamic adaptation to a specific host.

This projects main aim is to explore the mechanisms that underlie the broad host range. It is can be readily hypothesised that a key mechanism by which *P. capsici* can dynamically adapt to distinct hosts is through differential expression of genes. Therefore the specific aim of this project is to analyse this differential expression of genes in response to different host plants. And to characterise and identify those key genes that are differentially expressed during distinct host infection.

However, the genomic tractability of *P. capsici* is limited, and characterisation experiments can only give so much insight conducted with the current tools.

Additionally, transcriptomic studies of key lifecycle stages, notably those early stages of infection are hard to carry out, especially when looking for differential expression of genes. In light of these limitations, new tools for genomic editing and methodologies for transcriptomic *in vivo* studies are required. Therefore this project has also aimed to enable the deeper study of *P. capsici* through the development of such tools.

Chapter 2: Implementation of a TRAP-SEQ technology for the study of gene expression and translation in *P. capsici*

2.1 Introduction

Previous studies have highlighted the importance of the differential expression of genes in the dynamic adaption of a pathogen to a specific host and how this may be a mechanism that enables pathogenicity on multiple distinct host plants (Harris et al., 2016, Yang et al., 2018, Allan et al., 2019). To investigate the gene regulation that defines host adaptation, understanding of the early events that lead to the required changes is key. For this, an unbiased “omics” style approach to explore and define the earliest stages of infection is required. The project seeks to thereby identify key host preferential effectors and regulators, regulation patterns of host specific virulence and dynamic host adaption, in order to better understand the mechanisms that underlie broad host range in oomycetes.

To do this, transcriptomic analysis of the gene expression during leaf infection is necessary. However, *in vivo* transcriptomic analysis is limited due to the presence of host-derived mRNAs that dominate the samples and diminish signals in most if not all, high throughput gene profiling technologies. Broad, and deep data sets, describing pathogen gene expression, are hard to acquire due to the scarcity of pathogen mRNA in the early stages of infection. Isolation and enrichment of RNA, either prior or post sequencing, is therefore necessary. Therefore, to address the inherent limitation of the host-pathogen system, attempts were made to implement a novel TRAP (Translating Ribosomal Affinity Purification) followed by RNA sequencing (TRAP-SEQ) protocol in *P. capsici*. Thus in this context TRAP-SEQ enables the isolation of translating mRNA

of *P. capsici* from that of the material of the host plant, allowing the definition of early transcriptional changes that underpin dynamic adaptation and enable host invasion. This data alongside data from microarray experiments (Chapter 1) aids the definition of a suite of genes that are critical to distinct host virulence.

2.1.1 Translating Ribosome Affinity Purification – RNA sequencing

To analyse the transcriptome of *P. capsici* as described above there is a need to purify the mRNA of the pathogen from the excess of plant material, derived from an infected leaf. TRAP allows for the isolation of translating ribosome/ polysome complexes from a specific cell type, from a pool of total cell extract, in this case *P. capsici* material from infected leaf material (Figure 2.1). To do this in *P. capsici*, it was necessary to optimise and validate the TRAP protocol. TRAP involves cloning and epitope (FLAG) tagging of a ribosomal protein (RP). Epitope-tagged ribosomal units are then expressed in the target organism (Gracida and Calarco, 2017) or cell type (Bertin et al., 2015). Immuno-precipitation of intact ribosomal complexes can then be used to isolate translating mRNA from specific cell types or stages (Reynoso et al., 2015). The methodology was initially described by Heiman et al. (2008) and was used to specifically isolate the translating mRNA from a rare subset of mice neurons. Heiman et al. used the 60S ribosomal protein L10a tagged with eGFP for immunoprecipitation. More recent studies have isolated other cell types and from a diverse range of organism, from *X. laevis* (Yoon et al., 2012) (Watson et al., 2012) to zebrafish (Fang et al., 2013) (Tryon et al., 2013), *Arabidopsis thaliana* (Jiao and Meyerowitz, 2010) and *Drosophila* (Thomas et al., 2012, Huang et al., 2015). These studies have also used a variety of tagged large ribosomal protein subunits (RPL) and epitopes for

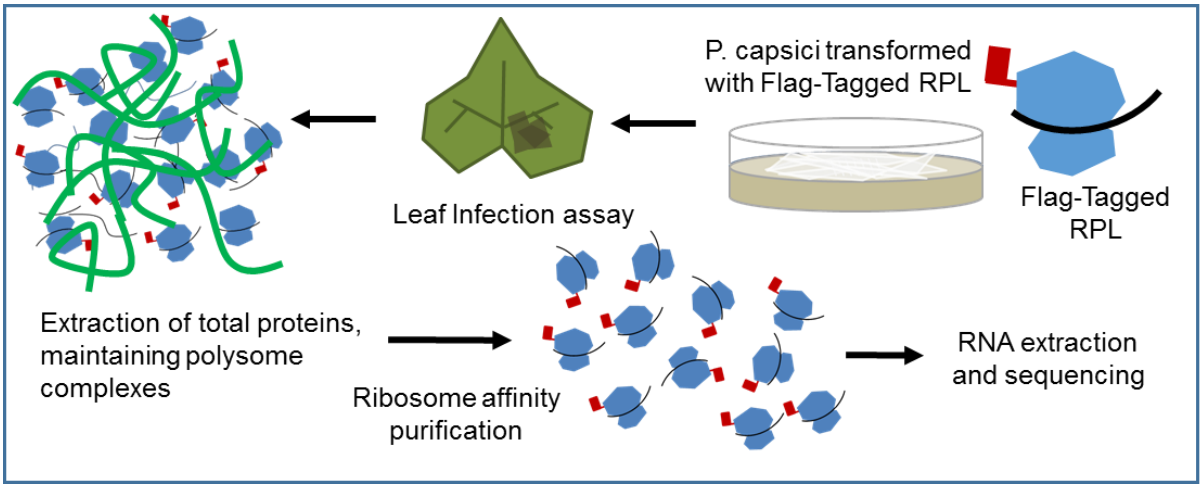


Figure 2.1: Schematic of TRAP-SEQ methodology.

This schematic shows the FLAG tagged ribonucleic complex, the FLAG tagged RPL of which was transformed and overexpressed in *P. capsici*. The strain of FLAG tagged RPL overexpressing *P. capsici* was then used to infect a host leaf. Following the progression of infection, total extract taken in a ribosomal maintenance media which preserves the integrity of the polysome. The total extract contains an excess of plant material in addition to *P. capsici* material which includes the intact polysomes. Immunoprecipitation of the ribonucleic complex using the FLAG tag, was used to isolate the *P. capsici* translating mRNA still attached to the intact polysome. The translating mRNA from these ploysomes was then isolated and used for RT-PCR or Sequencing.

immunoprecipitation. In Arabidopsis, RPL18 with both His-FLAG tags and a GFP tag were utilised, and in Mice a HA-tagged RPL22 has been used (Lyons et al., 2020). Both Drosophila and *C. elegans* both utilised GFP-tagged RPL10a (Thomas et al., 2012, Gracida and Calarco, 2017).

The aim of this project was to use the TRAP protocol, as described, to isolated *P. capsici* translating mRNA from infected plant tissue, and thus sequence the *P. capsici* translatome in the early stages of infection. It was hypothesised that by epitope-tagging and isolating ribosome-mRNA complexes solely from *P. capsici*, TRAP-SEQ will thus enable an increased breadth, and depth of sequencing by isolating only *P. capsici* genes. If true, a TRAP-SEQ approach will enable a more in-depth analysis of

changes in the *P. capsici* transcriptome during distinct host infection. By isolating and sequencing the translatome, TRAP-SEQ may also improve the correlation between transcript and protein levels, as only actively transcribed mRNAs are sequenced. Indeed, TRAP-SEQ has been shown to enable the isolation and sequencing of actively-translating mRNA derived from purified polysomes, which can give a more accurate view of protein production and content (de Sousa Abreu R, 2009).

The aim of this chapter was therefore to validate TRAP-SEQ methodology, novel to oomycetes and a host-pathogen bio-system, and to utilise it to examine the differences in the molecular mechanism of host adaptation to two distinct *P. capsici* hosts (tomato and cucumber). A strain of *P. capsici* overexpressing a FLAG tagged RPL was created, immunoprecipitation of the ribonucleic complex from *P. capsici* mycelial tissue grown *in vivo* was performed. Isolation of the FLAG tagged RPL22 was shown to be possible, however, purification of the translating mRNA from isolated FLAG tagged RPL22, presumed to be in a polysome complex, proved challenging. Thus this chapter also outlines details of the optimisation of the TRAP methodology from infected plant leaves, and the potential further work that may be necessary to successfully implement this technology in *P. capsici*.

2.2 Results

2.2.1 Identification of *Phytophthora capsici* ribosomal proteins for candidates to clone for ribosomal purification.

Within the *P. capsici*, genome annotation of RP is incomplete in places and erroneously redundant in others. Therefore, it was necessary to manually annotate proteins for use with TRAP-SEQ. The *P. capsici* genome (Lamour et al., 2012a) was searched and genes extracted using the gene ontology (GO) term “ribosome”

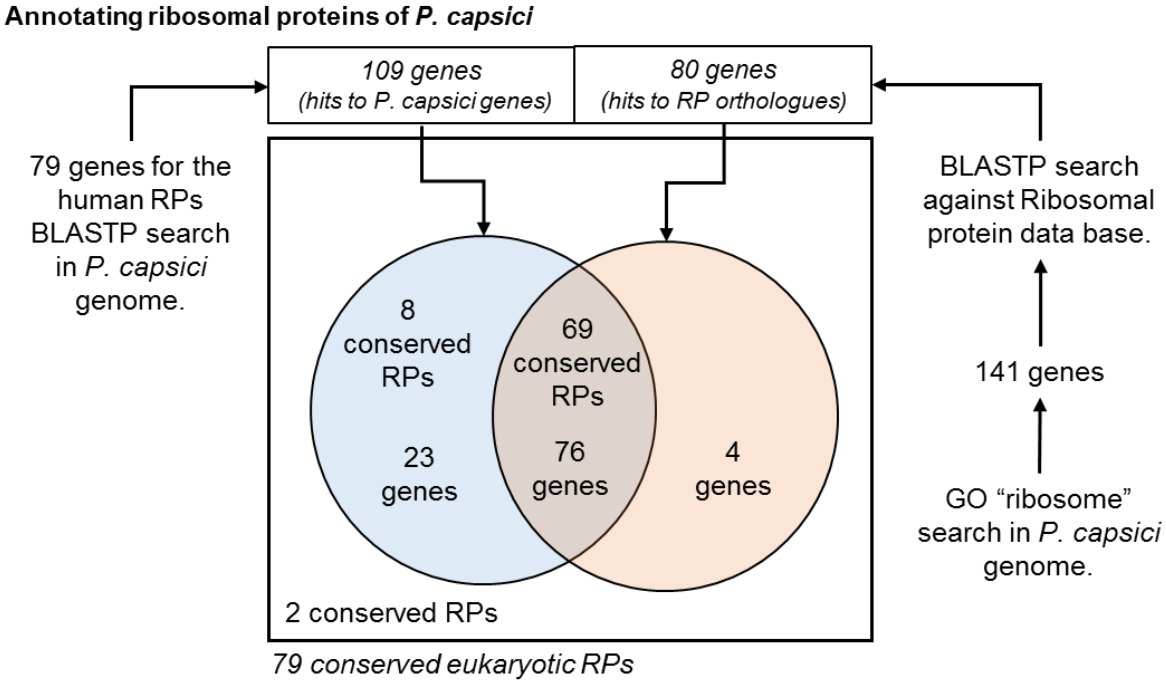


Figure 2.2: Schematic showing the RPL selection process.

Initially two different selection processes where used to conform ribosomal annotations on the *P. capsici* RPL. The 79 annotated human RPs were blast searched against the *P. capsici* genome giving 109 genes and then the 141 gene annotated with the GO term "ribosome" where blast searched against a ribosomal protein database. This gave 80 genes. The overlaps of these genes and which RP they were mapped to is shown in the pie chart. Overall 77 of the conserved eukaryotic RP where found. These were used when selecting the RPL for cloning a FLAG tagging for the TRAP methodology.

(GO:0005840). This identified 141 potential ribosomal protein genes, Sequences, corresponding to these candidates were then searched by BLASTP against a database of ribosomal protein sequences (Nakao et al., 2004). No hits were found for 34 genes and 27 matched mitochondrial ribosomes. These analyses left 80 genes assigned to 69 of the 79 conserved eukaryotic RP. To confirm these hits the human RP sequences were BLASTP searched against the *P. capsici* genome (Lamour et al., 2012a). Hits obtained in these analyses were compared to the original candidates. This resulted in 76 putative RP gene sequences found in both methods. The 76 genes

were the candidate genes used to pick the RP to be tagged, they covered 69 of the 79 conserved eukaryotic RP (Figure 2.2).

Within the literature no information seems to be available with regards to ribosomal protein selection for use in TRAP, although in all cases, the larger RPL subunit is used. RPL18 has been used multiple times in plants (Slane et al., 2014, Jiao and Meyerowitz, 2010), whereas RPL10A has been used in *C. elegans* (Gracida and Calarco, 2017) and RPL22 has been used in mice (Lyons et al., 2020). Therefore, the orthologues of RPL18, RPL10A and RPL22 (in this case two distinct orthologues were found PcRPL10Aa, and PcRPL10Ab) were used for cloning giving four FLAG-epitope-tagged ribosomal proteins: (Phyca11_9769: PcRPL18, Phyca11_15914: PcRPL10Aa, Phyca11_39189: PcRPL10Ab, Phyca11_19566: PcRPL22). BLASTP search of PcRPL10Ab against the *P. infestans* genome gave two results, one a putative uncharacterized protein and the other a 60S ribosomal protein: L10a-1 (PITG_22135T0), with which it has low homology. Additionally, the original BLASTP search against the ribosomal protein database gave poor homology, and the BLASTP of the human RP genes did not find this gene as it did with PcRPL18, PcRPL10Aa and PcRPL22. That being the case PcRPL10Ab was omitted from further consideration. Note that PcRPL18, PcRPL10Aa and PcRPL22 were also searched for using BLASTP against the *P. infestans* genome resulting in PITG_11099:60S ribosomal protein L18-2, PITG_22135: 60S ribosomal protein L10a-1 and PITG_01042T0: 60S ribosomal protein L22, respectively. Finally, these three candidates were searched in the four host Microarray data (Unpublished Huiteme Lab – Chapter 1) to check that these genes were not differentially expressed in any of the four hosts. Thus these RPLs were cloned, FLAG tagged and transformed into *P. capsici*. A modified Agrobacterium mediated transformation methodology was used (Vijn and Govers, 2003),

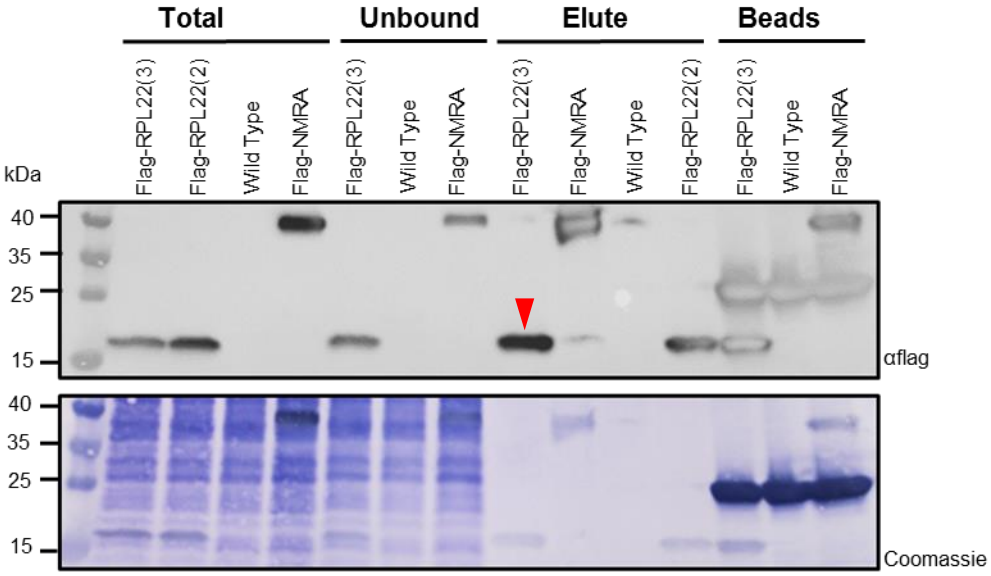


Figure 2.3: Western blot of TRAP pull down from Mycelia.

From mycelia grown on an agar plate with 2 FLAG-tagged RPL22 overexpressing strains (2) and (3), a Flag-NMRA overexpressing line as a positive control, and a wild type strain (WT) total protein extract was obtained. Total protein extraction was used as material for immunoprecipitation of the FLAG tagged ribosome. Here the total extract, unbound fraction, elute, and the remaining beads where run on and SDS-page gel. The gel was probed with anti-FLAG antibody, and visualised with a secondary antibody conjugated to HRP. Loading control was conducted using coomassie. Note that for the unbound fraction and the beads the strain Flag-RPL22(2) has been omitted. The red arrow indicates the protein band of Flag-RPL22 following elution.

unfortunately due to low transformation efficiency only transformants expressing FLAG tagged PcRPL22 were obtained.

2.2.2 Ribosomal Immunoprecipitation

Immunoblots were used to check the presence of FLAG tagged PcRPL22 and the effectivity of immunoprecipitation, the expected band (17 kDa) from total mycelia extract was present (Figure 2.3). Two transformants were tested for the presence of the FLAG tagged PcRPL22 (FlagRPL22(3) and FlagRPL22(2)). Immunoprecipitation using anti-FLAG magnetic beads following total extraction of protein using RMM media enable isolation of PcRPL22, however, this process is not 100% efficient as FLAG

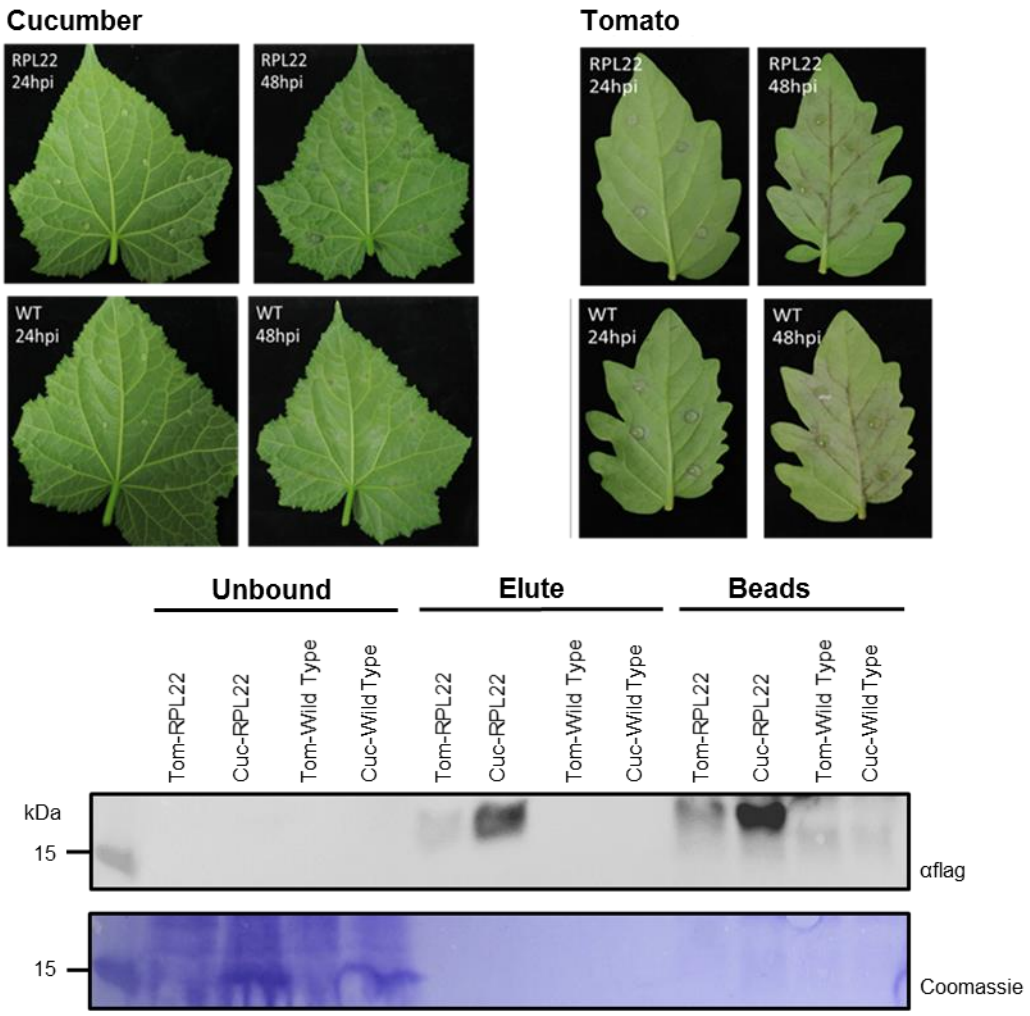


Figure 2.4: Leaf infection assay and Western blot of TRAP pull down from infected leaf.

P. capsici drop inoculated infected leaves from cucumber and tomato plants, picture taken at 24 hours post inoculation (HPI) and 48 hpi. For both host infection with FLAG-tagged RPL22 overexpressing strain and a wild type strain (WT). Total protein extraction from leaves, from both host plants and both strains was used as material for immunoprecipitation of the FLAG tagged ribosome. Unbound faction, elute and the remaining beads where run on a SDS page gel, and probed with anti-FLAG antibody, and visualised with a secondary antibody conjugated to HRP. Loading control was conducted using coomassie.

tagged protein remains in the unbound fraction and bound to the magnetic beads (Figure 2.3). However, immunoprecipitation was successful and a band for the FLAG tagged RPL22 can be seen in the eluted fraction for *P. capsici* strains FlagRPL22(3) and FlagRPL22(2).

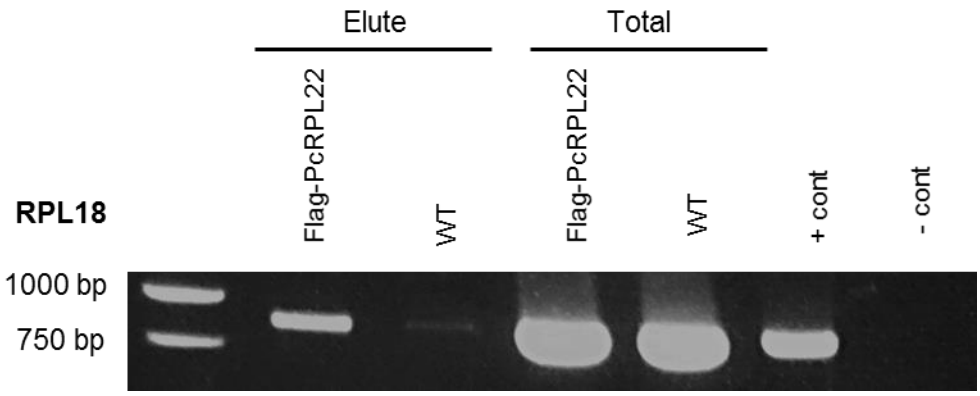


Figure 2.5: RT-PCR of control gene RPL18 form TRAP from mycelia.

From mycelia grown on a agar plate a FLAG-tagged RPL22 overexpressing strain and a wild type strain (WT) total extract was obtained. From this translating ribosomal complexes where purified and translating mRNA derived thereof. RT-PCR of this mRNA amplify the constituently expressed RPL, RPL18. RT-PCR of RNA from total extract was also conducted. Here the positive control is a plasmid containing RPL18 used as template DNA and the negative control is water used as the DNA template.

Two host plant (cucumber and tomato) were then infected with wild-type *P. capsici* (LT1534) alongside FlagPcRPL22 *P. capsici*, infection progressed the same (Figure 2.4) in each host when comparing the wild type strain (WT) to the FLagRPL22 overexpressing strain. The presence and successful immunoprecipitation from infection of cucumber and tomato were investigated with immunoblot. Immunoprecipitation was achieved and qualified by the observation of the expected band, although expression seems stronger in cucumber compared to tomato (Figure 1.4).

2.2.3 RNA from Ribosomal purification

FlagPcRPL22 *P. capsici* allows isolation of ribosomal proteins. To check for the presence of translating mRNA from *P. capsici* RT-PCR of ribosomal complexes purified from total mycelia extract was conducted. The presence of FLAG tagged ribosomal proteins allowed the amplification of the constitutively active PcRPL18 mRNA following ribosomal complex purification, whereas WT extract showed no

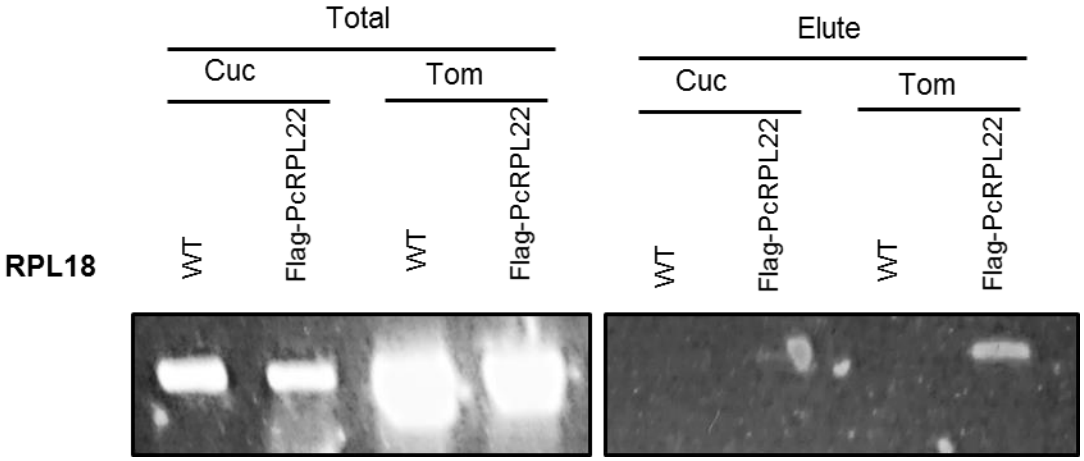


Figure 2.6: RT-PCR of control gene RPL18 form TRAP from leaf infection assay.

P. capsici drop leaf infection assay was conducted on leaves from cucumber and tomato plants, For both host infection with FLAG-tagged RPL22 overexpressing strain and a wild type strain (WT). Samples were taken at 24 hours post inoculation, total extraction from both host plant leaf infections and both strains was used as material for TRAP. mRNA was derived thereof. RT-PCR of this mRNA (Elute) was conducted to amplify the constituently expressed RPL, RPL18. RT-PCR of RNA from total extract was also conducted (Total).

amplification. However, it appears there is mRNA in total extract which would be expected from mycelia extraction (Figure 2.5). The experiment was repeated using infected leaves, here the presence of FlagPcRPL22, following RP purification, showed weak amplification of the housekeeping gene, almost none from infected cucumber leaves and low levels from infected tomato leaves. Again, greater quantities of mRNA appear to be gained from total extraction, and a strong band for the amplified PcRPL18 was observed (Figure 2.6). However, this result was not repeatable, and it proved challenging to obtain any mRNA from purified ribonucleic complexes from infected leaves.

2.2.4 Methodology optimisation

Aspects of this methodology were changed in order to gain reliable, good quality translating mRNA from infected leaf tissue. Efforts were made in three main areas to

optimise the protocol 1) the leaf infection assay itself, 2) the extraction and immunoprecipitation, and 3) RNA extraction from immunoprecipitated ribosomes/ polysomes. As it was possible to reliably isolate translating mRNA from mycelial tissue, the reason that unrepeatable, low quality and low quantity mRNA was derived from infected leaf tissue was perhaps due to the very small amount of *P. capsici* tissue present in the sample. Initial optimisation efforts centred on attempts to increase the amount of *P. capsici* tissue present in the infection assay. Changes in concentration and leaf inoculation methods were trialled. The initial methodology was to place a 10 µL drop of isolated zoospores, at $1x10^6$ spore per mL onto the abaxial side of the leaf. Development of the methodology consisted of, larger droplets up to 200 µL of zoospore suspension added to the leaf surface, and then a spray inoculation where the whole abaxial side of the leaf was sprayed with zoospore suspension with a misting bottle was also trailed. Finally, the leaf was placed abaxial side down into 20 mL of zoospore suspension for the first 2 hours of the infection assay. Whilst *P. capsici* was able to successfully colonise the leaf in all these inoculation techniques, they did not have any effect on the ability to pull out *P. capsici* translating ribosomes and successfully isolate mRNA thereof.

With the extraction technique there was little that could be changed, the buffer was well described by other methodologies, and is key for maintaining polysome integrity. The aspects that were changed were whether the tissue was flash frozen in liquid nitrogen prior to lysis, or if lysis was carried out in the buffer or without. Tissue lysis was also carried out using a mortar and pestle or using the TissueLyser II (QIAGEN). Attempts to increase the immunoprecipitation efficiency were also carried out by increasing the volume and number of elutions, and also the time the extract was given to bind and the temperature it was kept at during the binding process. However,

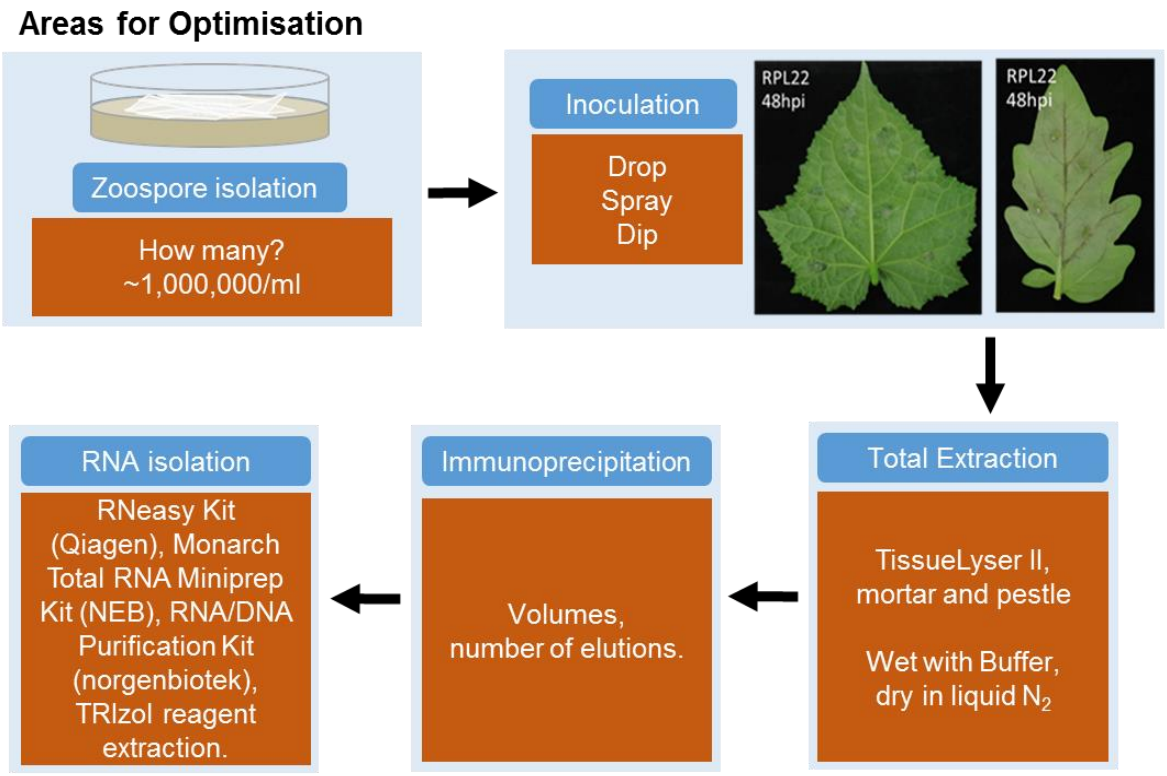


Figure 2.7: Schematic of Optimization.

Various stages of the TRAP protocol were changed in order to obtain translating mRNA from infected leaves. Each stage were subject to some change, from the isolation of zoospores for infection to the RNA isolation following TRAP.

binding and elution never reached 100% as FLAG tagged ribosome always remained in both the unbound fraction and attached to the bead. Finally, several different RNA extractions methodologies were used to try and extract RNA from the purified ribosomal complexes, if indeed there were complexes present. Three kits were used, RNeasy Mini Kit (Qiagen), Monarch Total RNA Miniprep Kit (NEB) RNA/DNA Purification Kit and (Norgenbiotek). In addition to the RNA isolation protocol using TRIzol RNA Isolation Reagents (Thermo Fisher). Again, none of these optimisations had any effect on the ability to isolate translating mRNA from *P. capsici* during leaf infection (Figure 2.7).

2.3 Methodology

2.3.1 Ribosomal protein identification

The *P. capsici* genome (Lamour et al., 2012a) and genes extracted using a search for the gene ontology term “ribosome” (GO:0005840). Potential ribosomal protein genes were then searched by BLASTP against a database of ribosomal proteins (http://ribosome.med.miyazaki-u.ac.jp), To confirm hits the human ribosomal proteins were BLASTP searched against the *P. capsici* genome (Lamour et al., 2012a). Following investigation of the literature, four candidate RPL were chosen, Phyca11_9769: PcRPL18, Phyca11_15914: PcRPL10Aa, Phyca11_39189: PcRPL10Ab, Phyca11_19566: PcRPL22. The candidate genes were then blasted against the *Phytophthora Infestans* genome to confirm homology and annotation. Expression of these three genes was also checked in *P. capsici* whole cell baseline, Microarray data and four hosts differentially expression microarray data to check for good basal expression but not differential expression induced by different host plants. This resulted in three candidate genes for cloning and use in the TRAP-SEQ methodology. Phyca11_9769: PcRPL18, Phyca11_15914: PcRPL10Aa, Phyca11_19566: PcRPL22.

2.3.2 Generation of RPL over expression vectors

Following a literature search for TRAP-SEQ protocols, three candidate RP were selected for use in vector generation and the TRAP-SEQ procedure: *P. capsici* Ribosomal protein L22 (PcRPL22), L18 (PcRPL18), and L10A (PcRPL10A). Primers were designed to N-terminally FLAG epitope tag these three PcRPLs for expression in *P. capsici* using the Agrobacterium vector pCB301TOR. The MCS of pCB301TOR contains an EcoRI site which was chosen for linearization prior to cloning in the

amplified PcRPL PCR products. The Infusion method of cloning (Takarabio) was used as per the manufactures instruction. Full-length CDS for the three candidate RP were cloned from total mRNA extractions. PCR amplicons were then transferred into the vector pCB301TOR, sequence fidelity was checked with sequencing and correct vectors were transferred into *Agrobacterium tumefaciens* for use in the transformation of *P. capsici.*

2.3.3 P. capsici Transformation

Agrobacterium mediated transformation of *Phytophthora* was modified from (Wu et al., 2016). Agrobacterium containing the desired plasmid, two days prior to transformation were spread onto new Lysogeny broth (LB) plates, containing the necessary selection antibiotics. To do this one large colony was dissolved in 50 µL of water and then spread evenly across the entire plate. In two days this results in a completely carpeted plate. A portion of these cells were then collected by pipetting up a down with 5 mL of *Agrobacterium* induction medium (IM). IM medium contains per 1 litre, 800 µL 1.25 M K_2HPO_4 pH 4.8, 1 mL 1% $CaCl_2$, 10 mL 0.01% $FeSO_4$, 40 mL 1 M MES buffer pH 5.5, 10 mL 50% glycerol, 2.5 mL 20% NH_4NO_3, 20 mL 20% glucose, 5 mL Microelements solution (containing 0.1% w/v of $ZnSO_4.7H_2O$, $MnSO_4.H_2O$, $CuSO_4.5H_2O$, $Na_2MoO_4.7H_2O$, and H_3BO_3), and 20 mL MN buffer (containing 3% w/v $MgSO_4.7H_2O$, 1.5% w/v NaCl) with the addition of 200 µM Acetosyringone). These were then agitated in the dark for 2 hours at room temperature on a shaker (60 rpm) to induce virulence gene expression. After 2 hours on the shaker, the optical density (OD) was examined, the Agrobacterium suspension was then diluted to 0.4 OD with IM medium. Zoospores were collected. *P. capsici* LT1534 was grown on 90 mm V8 agar plates in the dark at 25°C sealed with parafilm. This was followed by 2 days of growth in the light at 22°C without parafilm. Zoospores were collected by flooding the

90 mm plate with 10 mL of sterile distilled water at room temperature, then the mycelia growth was agitated with a sterile plate spreader and everything was transferred to a second *P. capsici* inoculated V8 agar plate. Continuing in this fashion, the 10 mL of water was used to flood and agitate three 90 mm plates recovered into a 50 mL falcon tube. The sporangia suspension was left for 30 minutes with the lid removed on a lightbox to induce zoospore release. The suspension was then filtered through one layer of Miracloth to remove mycelia and any agar chunks in the suspension, this process yielded approximately 5 mL of ≥ 1 x 10^6 spores mL-1 zoospore suspension. Equal volumes of Agrobacterium and zoospore suspension (2-5 mL each) were added together and gently swirled. The mixture was then incubated for 2 hours in the dark at room temperature. Then 500 µL of the mixture was placed on top of a 5x5 cm piece of sterile Hybond N+ membrane, which had been placed in a 90 mm solid IM plate, and dried for 10 minutes. IM solid is IM medium as previously described with 1.5% w/v agar. The plates were then incubated in the dark for 2 days at room temperature. After 2 days the membrane should be coated in fluffy mycelia, it is then transferred upside down to a Plich medium. Modified Plich media contains per 1 litre, 0.5 g KH_2PO_4, 0.25 g $MgSO_4.7H_2O$, 1 g Asparagine, 1 mg Thiamine, 0.5 g Yeast extract, 10 mg β-sitosterol, 25 g Glucose, and 15 g Agar with 50 µg/mL G418 and 200 µM cefotaxime. It is then incubated in the dark for a further 3 days at room temperature. The membrane is then removed and the plate checked at 1-3 days for G418 resistant colonies. These are transferred to V8 medium plates containing 50 µg/mL G418 and 200 µM cefotaxime and checked for expression of transgenes.

2.3.4 Immunoblotting to check the expression of tagged RPL

P. capsici mycelia mats were grown in 10 mm plates with 25 mL pea broth, following 3 days of growth mycelia mats were harvested and total protein extracted. For the purposes of total protein, isolation mycelia was harvested and lysis buffer GTEN (containing: 10% v/v Glycerol, 25 mM Tris buffer pH 7.5, 1 mM EDTA, 150 mM NaCl, 10 mM DTT, 2% w/v PVPP, 0.1% v/v tween, and 1 mM PMSF) was added. Tissue was disrupted with a TissueLyser II (Qiagen) for 1 minute at 30 Hz then adapter flip and shake repeated. The resulting lysate is spun at 14,000 xg at 4°C for 10 minutes. The supernatant is taken the protein suspension was then diluted in equal volumes 2x SDS page buffer, boiled for 5 minutes at 95°C, and loaded onto Mini-PROTEAN TGX precast gel 4-20% (Bio-rad). Separated proteins were then transferred onto a nitrocellulose membrane. The membrane was subsequently blocked with 10% dried skimmed milk (Marvel). Membranes were probed with primary anti-FLAG (1:5000) for 60 minutes at room temperature. Subsequently, they were incubated with horseradish peroxidase labelled secondary antibodies (1:10,000) (LI-COR) and analysed.

2.3.5 Optimization of TRAP

The protocol was adapted from Mauricio A. Reynoso et al. (2015). The TRAP-RNA extraction process requires a FLAG tagged RP to be overexpressed in *P. capsici*. This allows isolation of translating mRNAs trapped in the ribosomal complex. For isolation of translating mRNA either *P. capsici* overexpressing FlagPcRPL22 was grown for 3 days in pea broth at 25°C in the dark, or zoospores were isolated as above, diluted to a concentration $1x10^6$ zoospores/mL, host leaves (tomato and cucumber) were completely submerged in zoospore solution the left in a humid environment at room temperature and the infection left to progress for 24 hours. Then total leaf extract was

obtained using ribosomal maintenance media (RMM, containing: 100 mM Tris–HCl (pH 9.0), 200 mM KCl, 25 mM EGTA (pH 8.0), 36 mM $MgCl_2$, 1% Brij L23, 1% Triton X-100, 1% IGEPAL CA-630, 1% TWEEN 20, 2% polyoxyethylene (10) tridecyl ether, 1% deoxycholic acid, 5 mM DTT, 1 mM PMSF, 50 µg/mL cycloheximide, 50 µg/mL chloramphenicol, 40 U/mL RiboLock RNase Inhibitor). 700 µl of RMM was added to each sample then shaken with a TissueLyser II (Qiagen) for 1 minute at 30 Hz then adapter flip and shake repeated. Lysed tissue was spun at full speed for 10 minutes at 4°C, and whole lysate supernatant collected. Total lysate was then used for immuno-precipitation of FLAG tagged ribosomal complexes. 30 µL of M2 anti-FLAG magnetic beads were added to 300 µL of total lysate, then incubated for 1 hour at 4°C. Following 5x washed with RMM, the sample was eluted twice with 150 µL of 250 µg/mL 5xFLAG peptide for a total of 300 uL eluted ribosomal complexes. Elute was either frozen at 80 °C or RNA extracted directly.

2.3.6 RNA Extraction

Multiple extraction methods of RNA extraction were used and quality compared. Initially, RNeasy (Qiagen) was used to extract RNA, due to the potential loss of short mRNAs the method was changed to use TRIzol reagent extraction. Monarch Total RNA Miniprep Kit (NEB), and RNA/DNA Purification Kit and (norgenbiotek) were also trailed for their effectivity. For the Kits extraction was conducted as per the manufactures instruction. For extraction using TRIzol reagent, 300 µL of elute was added to 800 µL TRIzol reagent. This was then vortexed for 30 seconds and then left at room temperature for 5 minutes to allow the dissociation of the mRNA from the ribonucleic complexes. 200 µL chloroform was then added. This was then shaken well by hand until uniform and cloudy, and centrifuged from 15 minutes at 12,000 xg at 4°C. This step allows the separation of distinct phases, a bottom organic phase

containing proteins, lipids and other cell debris, a middle interphase consisting of a cloudy film of DNA, and the clear upper phase which contains the RNA. The upper phase was transferred into a fresh RNase free 2 mL Eppendorf and 500 µL of isopropanol added to precipitate the RNA, this was inverted gently and then either left at room temperature at 10 minutes or a -20°C overnight. This was then centrifuged for 5 minutes at 12,000 xg at 4°C. The RNA pellet was then washed with 1 mL of ethanol and then spun for 5 minutes at 7,500 xg at 4°C. The pellet was then resuspended in 50 µL. The quantity of RNA was then estimated by nanodrop and the presence of mRNA checked with RT-PCR.

2.3.7 RT-PCR

To check for the presence of mRNA, RT-PCR from the constituently active RPL18 gene was conducted. For cDNA synthesis, RQ1 RNase-Free DNase (Promega) treatment immediately followed by Superscript reverse transcriptase III with Oligo(dT20) was used (Invitrogen) as per the manufactures instruction. PCR of RPL18 was conducted using GoTaq® G2 Flexi DNA Polymerase (Promega) as per the manufacturer's instructions.

2.4 Discussion

2.4.1 Aims and outcome

The aim of this research was to develop a technology novel to oomycete research, TRAP-SEQ. If successful this would allow the isolation of *P. capsici*, or other pathogen material directly from infected plant tissue, material that could then be used in transcriptomic/translatomic analysis. This would open up a recalcitrant but key stage

in the *P. capsici* / pathogen lifecycle that is potentially understudied due to the difficulty of access for such omics style experiments.

It was shown to be possible to use TRAP to isolate, what can be presumed to be, translating mRNA from mycelial tissue, although, it was not possible to reliably isolate any translating mRNA from the pathogen during plant leaf infection. However, TRAP-SEQ used in *P. capsici in vitro* represents a technique by which isolating translating mRNA as appose to the steady state RNA transcripts can be purified from *P. capsici* mycelia. Sequencing translating mRNA would be a way of both validating RNA sequencing experiments, and analysis what resemblance steady sate RNA transcripts and translating RNA have to each other. This could be carried out in mycelia, and perhaps other time points where plant tissue is not present, sporangia and germinating cysts for example. Although it has not been shown here, the TRAP protocol would likely be functional with those time points. One weakness of this study is that these experiments were never carried out and the translating mRNA from mycelia was never sequenced.

2.4.2 Potential Problems and Further optimisation

The TRAP-SEQ methodology has been used in several organisms mRNA from distinct cell types or developmental stages. Although TRAP-SEQ has been applied in plants (specifically in *Arabidopsis*), it has never been used in a pathosystem. The technology, therefore, has immense potential in host-microbe model systems, as we know very little about the early stages of pathogen infection. The ability to isolate and enrich pathogen material from a host opens up a clear route to understanding cellular processes at the very early stages of infection. For this reason, further optimisation of

the protocol is worthwhile. In terms of optimisation of TRAP-SEQ during plant infection, it was originally assumed that the inability to isolate translating mRNA was because of low levels of *P. capsici* material in the extraction process. Although, strong protein bands for the tagged-RPL22 were detected, both from total extraction and following ribosomal pulldown. In order to increase the amount of *P. capsici* material present in the leaf, multiple changes to the infection assay where trialled. Nonetheless, optimisations did not result in any improvement. It may be that the presence of host material in the pull down affected the isolation of mRNA, it is possible that something about the presence of the plant material in the extract has an effect on either the viability of the mRNA or perhaps the integrity of the polysome complex. Tomato and cucumber leaves especially have a multitude of phenolic and antioxidant compounds (Silva-Beltrán et al., 2015, Yunusa et al., 2018), and that some of these metabolites interfere with RNA or ribonucleic complex integrity. This would explain why the pulldown was possible in mycelia but not from infected plants. However, it has been shown that the TRAP methodology is functional in *Arabidopsis thaliana*. A possible test to examine if the methodology is viable from infected leaves, would be to infected *Arabidopsis* (which although is not a true host to *P. capsici* can be infected under certain conditions) and see if translating mRNA can be successfully isolated from that system. This would show whether the host plant material itself is interfering with the TRAP. Further attempts to implement this technology would also be to try other RPLs, RPL10A was cloned and is a popular RPL used with this technology. However, although attempts to transform FlagRPL10A into *P. capsici* were made, no transformants where produced. It would also be conceivable to use other tags for the immunoprecipitation of the ribonucleic complex, His tags and eGFP are also commonly described in the literature (Gracida and Calarco, 2017, Watson et al., 2012).

2.4.3 Conclusion

The aim of this novel technology was to study the dynamic adaption of *P. capsici* to host plants whilst infecting a host leaf. Although the TRAP-SEQ methodology was implemented in *P. capsici in vitro*, currently it does not work with staring material from infected plant leaves. It is not yet clear why this is the case. However, although attempts were made to optimise the methodology, alternative techniques to study the same life-cycle stages are possible. Although TRAP-SEQ would have been a great boon to the study of plant pathogens during infection, a great deal of further work may be necessary to successfully implement this technology and easier alternatives could be pursued and have been explored in the following chapters.

Chapter 3: Transcriptomic analysis of *P. capsici* during an *in vivo* model of infection of two distinct hosts

3.1 Introduction

Phytophthora capsici is a global and devastating oomycete plant pathogen that threatens a wide variety of economically important food crops across the planet. The size of this threat is partially due to the broad host range of *P. capsici*, capable of causing disease on many commercially important crops, particularly those in the *Solanaceae* and *Cucurbitaceae* families, in total 40 different crop plant species have been reported as hosts (Drenth and Guest, 2004). However, our understanding of what mechanisms enable *P. capsici* pathogenicity on this diverse range of crops is very limited.

The majority of studies trying to understand the pathogenicity of key plant pathogens usually involve a single pathogen-host system (Näpflin et al., 2019). With an understanding of the plant immune system and pathogen virulence strategies, it can be confidently stated that the pathogen has to run a very specific gauntlet to be able to cause disease on a host plant: avoidance and suppression of the plant immune system, resistance to innate toxins and reactive oxygen species (ROS), intimate integration and growth within the host, and incorporation of the metabolic environment, all represent idiosyncratic challenges specific to each host. Therefore, the dynamic adaption of a pathogen to a host plant can be presumed to be very distinctive, and different host plants, especially those from different families, can represent very different environments for pathogen growth. The current conception of the key determinants of host specificity and host range is limited: it is not fully understood how

a pathogen can infect multiple host plants. Recent research that has been conducted examining host specificity and host range shows two main trends in pathogens with broad host ranges. Firstly, they have dynamic genomes, prone to expansion and/or mutation, and can create vast genetic diversity relatively quickly (Ma et al., 2010, Raffaele et al., 2010, Hacquard et al., 2013). Secondly, they display differential expression of genes upon different host infection (Yang et al., 2018, Harris et al., 2016, Allan et al., 2019), presumably a display of dynamic adaption to that specific host or environment.

It is already well documented that *P. capsici* is genetically diverse and has a remarkably pliable genome (Barchenger et al., 2018, Lamour et al., 2012a, Hu et al., 2020). Additionally, it is reasonable to hypothesise that differential gene expression acts as the driver of the dynamic adaption of *P. capsici* to its many hosts. However, to date, no studies have been conducted examining the pathogenicity differences of *P. capsici* on different host plants. Understanding the key mechanisms of dynamic adaption of *P. capsici*, especially in terms of its devastating broad host range is, therefore key to tackling the disease. It could also be hypothesised that early time points in the plant infection are key to host perception and therefore the fate decisions that yield the distinct virulence strategies in question. Therefore, transcriptomic studies of different host plant infections, especially at early time points, would grant great insight into the mechanics of host perception and the dynamic adaptions of *P. capsici*.

To further understand the differences in the virulence strategies, and the mechanisms that drive host-specific dynamic adaption, detailed transcriptomic studies are key. However, due to limitations discussed in previous chapters, it is not yet possible to conduct deep and broad transcriptomic analysis of *P. capsici* during the early stages of plant infection. No methodology yet exists to isolate and sequence the transcriptome

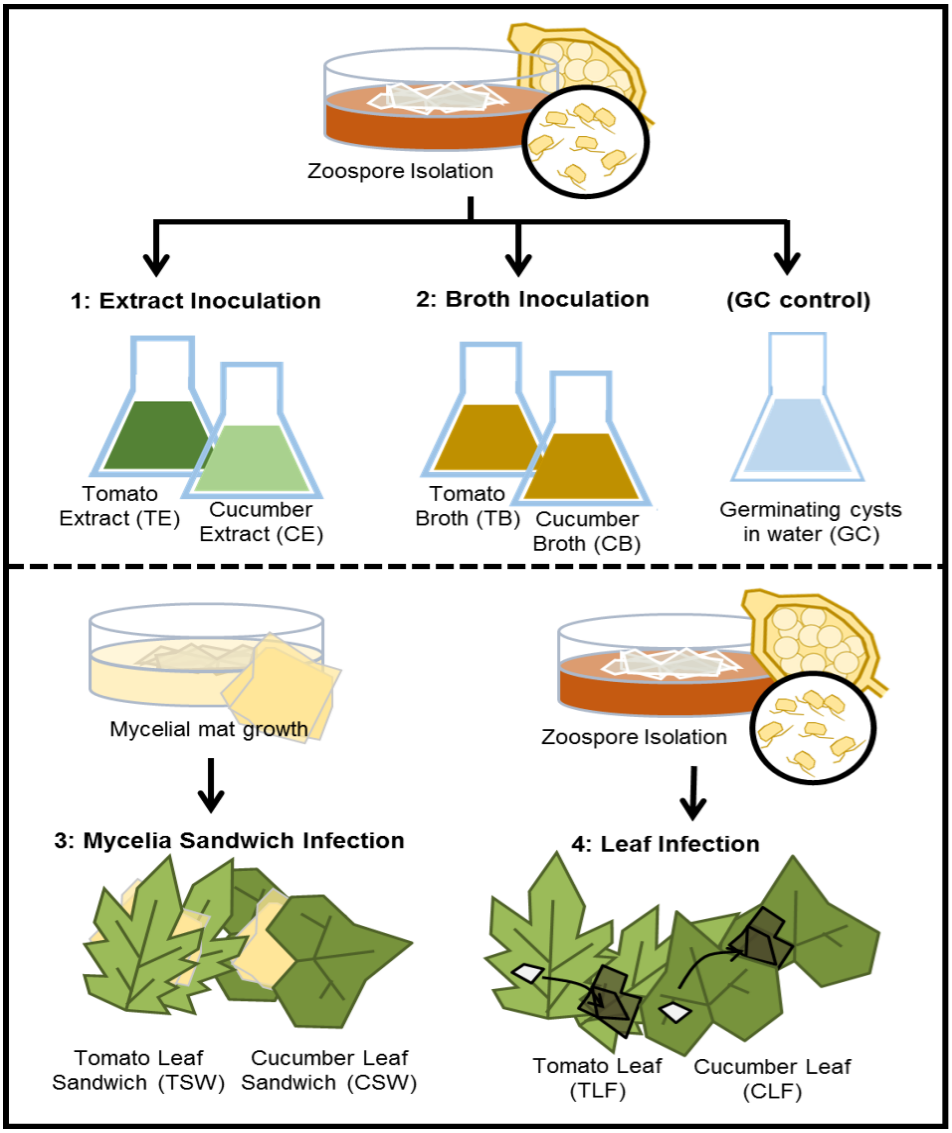


Figure 3.1: Schematic of leaf infection assay and zoospore inoculation assay methodology.

P. capsici was grown on V8 agar plates, it was induced to sporulate and the spores collected. From these sporangia zoospores were isolated and used for three different infection assays. 1) extract inoculation in both tomato (TE) and cucumber extract (CE), 2) broth inoculation in both tomato (TB) and cucumber broth (CB), also included here is the germinating cysts (GC) control, zoospores used to inoculate water. Zoospore where also used to conducted 4) leaf infection assays in both tomato (TLF) and cucumber leaves (CLF). In addition mycelia was grown in liquid pea broth to form mycelia mats, which where then used sandwiched between to leaves to conduct the 3) mycelial sandwich infection assay, conducted with both tomato (TSW) and cucumber leaves (CSW). Samples from two these infection assays were harvested at 2, 4 and 8 hours post infection (hpi). From all these samples total RNA was extracted and used for transcriptomic studies or qRT-PCR. Note this numbering is used to refer to these methodologies throughout the text.

of *P. capsici* whilst infecting plant tissue. Therefore, alternative methodologies were explored that do not require distinguishing *P. capsici* sequence reads from excessive and spurious host reads post-sequencing. Initially, three potential *in vitro* methodologies were trialled and assessed to identify which method resulted in gene expression patterns most similar to those observed in infected tissues.

The three methodologies were designed as *in vitro* models of infected plants, they comprise of 1) zoospores used to inoculate plant extract, 2) zoospore used to inoculate plant broth, and 3) mycelia leaf sandwiches (Figure 3.1). Briefly; zoospores used to inoculate plants broths or extracts comprised isolated zoospore suspension, mixed with either plant broth or plant extract and left to germinate and grow. What is referred to as plant broth in the chapter consisted of plant leaf tissue in water, autoclaved, and then filtered. Plant extract consisted of plant leaf tissue frozen in liquid nitrogen, ground into a fine powder, suspended in water and then filtered. Mycelia leaf sandwiches were comprised of two host plant leaves placed either side of a square of *P. capsici* mycelial tissue with the abaxial side in contact with the mycelia. These methodologies have the potential to be used with any host or non-host plants to examine both common and differential gene expression induced by that particular plant. This study allowed us to assess two host plants. Previous experiments had shown that from the four host plants available, cucumber and tomato were the two most different in terms of inducing differential gene expression in *P. capsici* during infection. Cucumber and tomato are members of the two plant families that make up the majority of the host plants of *P. capsici*, *Cucurbitaceae* and *Solanaceae* respectively. They were, therefore, two ideal host candidates to examine the mechanisms of dynamic adaption that drive broad host range. Understanding these two patho-systems has real-world applicability as they are both grown commercially and suffer from *P. capsici* caused crop losses.

Following examination of differential gene expression in all three *in vitro* systems, zoospores germinated and cultures in plant extract were found to be the most similar to infected plant leaves (Figure 3.2, Figure 3.3). This model system was therefore taken forward and total transcriptomic analysis by RNA sequencing was conducted (**Error! Reference source not found.**).

This study has found host dependent changes in the profile of *P. capsici* transcriptome *in vitro*. In early time points, cucumber extract induces an up-regulation of many genes not found to be induced by tomato extract. On the contrary, *P. capsici* response in tomato is surprisingly mostly characterised by a down-regulation event. These differentially expressed genes have been identified at early time points and were found to be largely integral components of the membrane with roles in oxidation/reduction, transmembrane transport, and phosphorylation. A large amount of differentially expressed effectors were also identified, they largely shared a similar expression profile for the other differential expressed genes, that is, induced in response to cucumber and down-regulation in response to tomato. These genes require characterisation and validation but may represent a bounteous list of genes with putative host specific roles. This understanding could be used to inform further mechanistic studies of host perception and dynamic adaption of *P. capsici* and other broad host range oomycete pathogens.

3.2 Results

3.2.1 Validation of in vitro methodologies by qRT-PCR of known genes

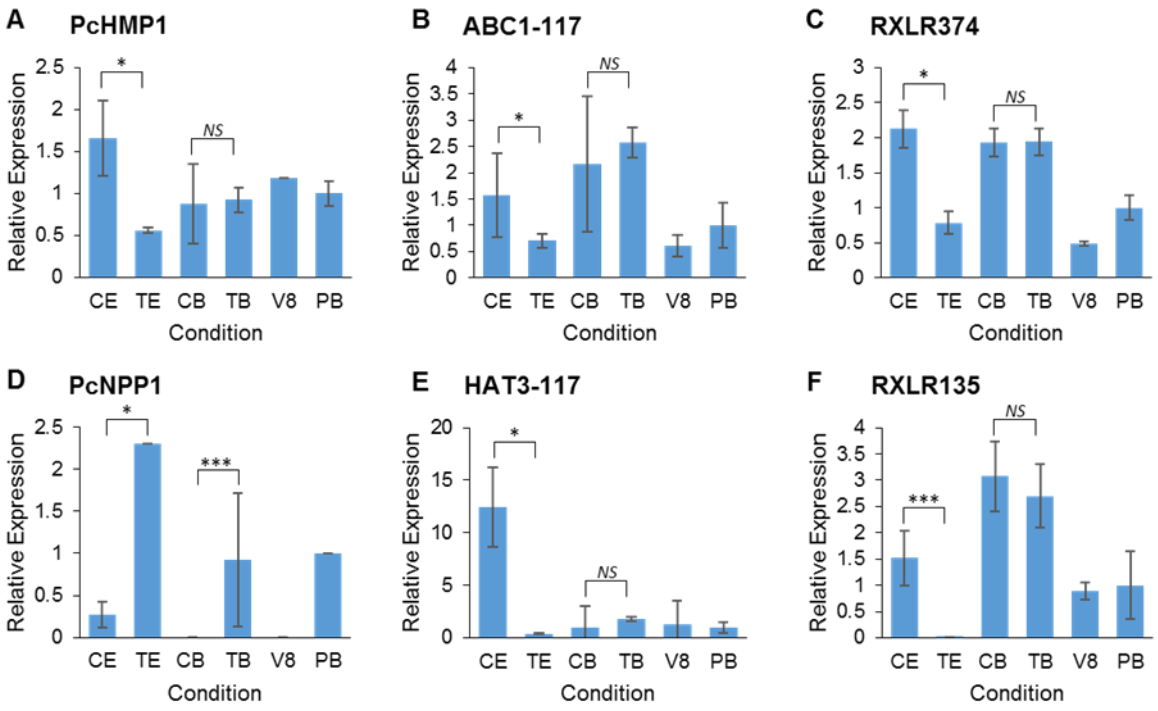


Figure 3.2: qRT-PCR expression of key gene in Extract and Broth inoculation.

Total RNA was obtained from 1) extract inoculation assay of tomato (TE) cucumber extract (CE), and from 2) broth inoculation assay of tomato (TB) and cucumber broth (CB). RNA was also isolated from inoculated V8 juice and pea broth (PB). qRT-PCR was used to analysis the expression of six genes in these six conditions. The six genes are PcHMP1 (A), ABC1-117 (B), RXLR374 (C), PcNPP1 (D), HAT3-117 (E), and RXLR135 (F). Expression relative to the house keeping gene Tubulin is shown and was calculated using the ΔΔCT method. Expression is shown relative to PB. Error bars show the SEM of at least 3 technical replicates. Statistical significance is shown for comparison between CE and TE, and CB and TB. Two tailed unequal variance student t-test was conducted, significant bonferroni adjusted P values are shown. Adjusted P values of <0.05 = *, <0.01 = **, <0.001 = ***, >0.05 = non-significant (NS). Note that Leaf infection assay and

The novel *in vitro* methodologies (1.Extract inoculation, 2.Broth inoculation, and 3.Mycelia sandwiches) of *P. capsici* culture that were trialled for potential use with transcriptomic analysis were conducted, and samples were taken. From these samples, expression data of genes that were known to have host dependent differential expression (DE) (unpublished Huitema Lab) were assessed through qRT-PCR. Of the three potential *in vitro* methodologies, only zoospore inoculation of Extracts and Broths were initially conducted. For these experiments two host plants

were used (tomato and cucumber), therefore a total of 4 culture conditions are possible Tomato Broth (TB), Tomato Extract (TE), Cucumber Broth (CB) and Cucumber Extract (CE). In addition zoospore inoculation of V8 juice (V8) and Pea broth (PB) were included. In the initial experiments cucumber leaf infection (CLF) and tomato leaf infection (TLF) were not included.

The expression of six genes was examined, two of these were marker genes PcHMP1 (Phyca11-506620) and PcNPP1 (Phyca11_11951) markers of biotrophy and necrotrophy respectively. A further four genes (ABC1-117/ Phyca11_512023, HAT3-117/ Phyca11_512020, RXLR374/ Phyca11_13953 and, RXLR135/ Phyca11_534478), which were all shown to be host dependently differentially expressed were tested. Specifically, these four genes were found to be up-regulated in CLF at 16 hours post-inoculation (hpi) with respect to 3 other *Solanaceae* host plants (unpublished Huitema Lab – Chapter 1). In all four of the genes tested the same host dependent differential expression, that being up regulation in cucumber based conditions, was observed if zoospores were used to inoculate the host extract (TE compared with CE) (Figure 3.2). Therefore, in *P. capsici* cultured in plant extract, genes known to be up-regulated in CLF are also specifically induced in CE and not induced in TE. It is also the case that the expression in CE is always higher than the two control conditions: zoospore used to inoculate V8 media and Pea Broth (PB). However, it was observed that culturing in host broths was not able to induce the same host dependent differential expression pattern as expected. CE presented higher expression of the marker gene PcHMP1 compared to TE and the other 4 culture conditions. However, the marker gene PcNPP1 was only induced in TE and had low expression in the other 5 conditions, especially CE, CB and V8 (Figure 3.2).

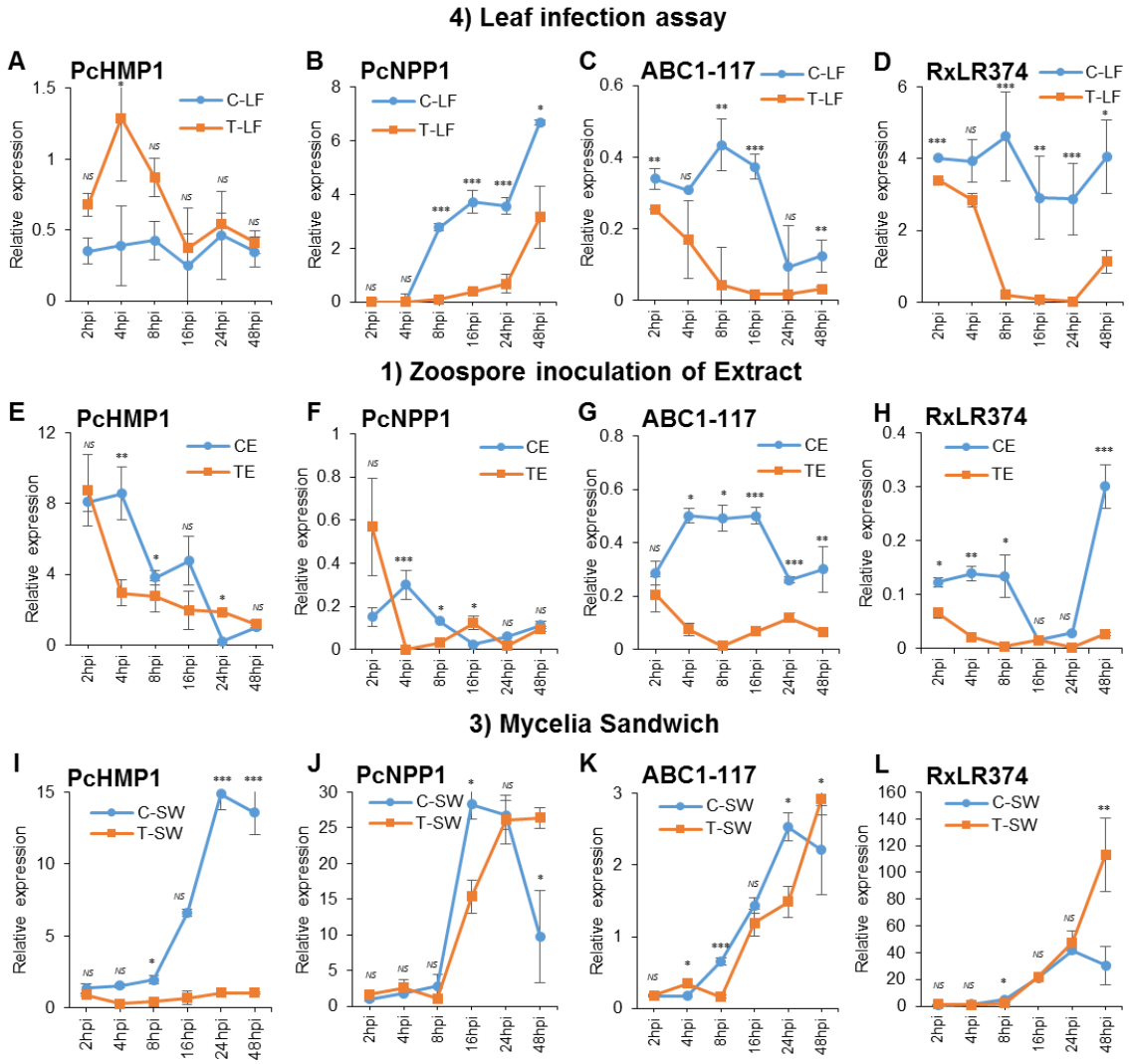


Figure 3.3: qRT-PCR expression of key gene in Extract inoculation, Leaf infection assay and, Mycelia Sandwich across a time course 2hpi – 48hpi.

Total RNA was obtained from leaf infection assays of both cumber (CLF) and tomato (TLF) (A-D), inoculation assay of cucumber (CE) tomato extract (TE) (E-H), and Mycelia sandwich infection assays of both cumber (CSW) and tomato (TSW) (I-L). For all six conditions RNA was collected at time points 2 hpi, 4 hpi, 8 hpi, 16 hpi, 24 hpi and 48 hpi. qRT-PCR was used to analysis the expression of 4 genes in these six condition in all time points mentioned. The four genes are PcHMP1 (A, E, I), PcNPP1 (B, F, J), ABC1-117 (C, G, K) and RXLR374 (D, H, L). Expression relative to the house keeping gene Tubulin is shown and was calculated using the ΔΔCT method. Expression is shown relative to GC 2 hpi samples which are not displayed in the graphs but are set to 1. Error bars show the SEM of at least 3 technical replicates. Statistical significance is shown for comparison between CE and TE at each time point. Two tailed unequal variance student t-test was conducted, significant bonferroni adjusted P values are shown. Adjusted P values of <0.05 = *, <0.01 = **, <0.001 = ***, >0.05 = non-significant (NS). Note that inoculation of broths has not been included here, and leaf infection assays have.

Following these experiments, time course data for both host extract inoculation and leaf infection assay was collected. In addition, a time-course experiment for the third *in vitro* methodology: mycelial sandwiches (Tomato leaf sandwich – TSW and Cucumber leaf sandwich – CSW) was conducted and samples collected. qRT-PCR real leaf infections. For these experiments, four genes were tested. The two marker genes PcHMP1 and PcNMP1 as mentioned previously, and two genes that were shown to be induced in cucumber leaf infection at 16 hpi (ABC1-117 and RXLR374) (Figure 3.3). For ABC1-117 and RXLR374, the host dependent differential gene expression in TE and CE was the same as the differential expression found in the leaf infection assay. However, this was not observed with the mycelial sandwiches (TSW and CSW), as there was no difference in gene expression between hosts, with expression seemingly increasing as the infection progresses in both TSW and CSW (Figure 3.3). In terms of the marker genes, PcHMP1 expression is similar when comparing leaf infection assay to zoospores in host extract, both having reductions in expression as the experiment progresses, with little host dependent change in expression. PcNPP1 expression, however, does not follow similar trends in CE, TE or leaf infection. PcNPP1 shows an increase in expression during the latter stages of infection in leaf, however in extract the expression of PcNPP1 is very low. Mycelial sandwiches show a completely different expression profile of the two marker genes tested, with increasing expression as the infection progresses (Figure 3.3). It was therefore decided that host extracts would be used for transcriptomic analysis of the effect that the host has on *P. capsici* during the early stages of infection.

3.2.2 RNA extraction and quality control

For Transcriptomics high-quality RNA was essential. From hosts extract inoculation assays, the initial experimental design resulted in poor quality RNA, this was characterised by a low 260/280 ratio (<2) and low 260/230 ratio (<1.7) as measured by nanodrop. Gel electrophoresis of these samples showed a smear, characteristic of degraded RNA. Multiple optimisation steps in the methodology for RNA extraction were trialled.

Following optimisation, a new methodology for growth in an extract that allowed removal of the plant extract before lysis of the cells was developed. This allowed extraction of better-quality RNA, characterized by high 260/280 and 260/230 ratios, >2 and >1.7 respectively (Table 3.1). When analysed with gel electrophoresis, the absence of smear indicated that RNA was not degraded. It could be suggested that an excess of plant extract during the lysis phase, and perhaps persisting through the extraction process was causing RNA degradation and low purity. Removal of the extract before extraction is therefore essential to gain good quality RNA from zoospore inoculation experiments.

To quantify degradation and therefore the integrity of the RNA, the high purity RNA was analysed by Tapestation, giving RNA integrity numbers (RIN) that should be greater than seven for human RNA samples. However plant RNA is known to contain many small RNAs compared to human-derived RNA and therefore gives a lower RIN, *Phytophthora* species are similar to plants in this way, that they contain a lot of short mRNA and therefore a lower RIN than expected (Die and Román, 2012). Indeed electrophoresis showed good integrity, especially in samples derived from the control

Table 3.1: RNA purity, quantification and integrity.

Extract inoculation of Tomato extract (here: Tom) and cucumber extract (here: Cuc) and samples taken at 2 hpi, 4 hpi, 8 hpi and 16 hpi. In addition germinating cysts at 2hpi (GC) were also sampled. Four biological replicates for each host extract and time point were obtained and RNA extracted. RNA purity and an estimation of concentration was obtained from Nanodrop analysis, purity is shown by the 260/280 and 260/230 ratios and an estimate concentration is shown here in ng/µL. Quibit analysis was used further quantify the quantity of RNA, here shown both in total amount (µg), and concentration (ng/µL). RNA integrity numbers (RIN) were obtained from Tapestation analysis and shown here.

			Nanodrop			Quibit		Tapestation
Number	Name	Bilogical Rep	ng/ul	260/280	260/230	ug	ng/ul	RIN
1	Tom 2 hpi	1	55.9	2.22	2.71	2519.2	53.6	3.4
2	Tom 4 hpi	1	71.7	2.2	2.45	2462.8	52.4	2.8
3	Tom 8 hpi	1	69	2.14	1.92	1005.8	21.4	2.6
4	Tom 16 hpi	1	412.2	2.15	2.24	13160	280	2.6
5	Cuc 2 hpi	1	45.9	2.2	2.22	1192.5	26.5	4.1
6	Cuc 4 hpi	1	109.5	2.13	2.09	2171.4	46.2	4.2
7	Cuc 8 hpi	1	256	2.16	2.15	8413	179	2.7
8	Cuc 16 hpi	1	116.3	2.11	2.14	8554	182	1
9	Tom 2 hpi	2	23.8	2.27	1.49	1184.4	25.2	3.1
10	Tom 4 hpi	2	100.3	2.17	2.22	3816.4	81.2	2.9
11	Tom 8 hpi	2	121.5	2.22	2.46	5123	109	2.4
12	Tom 16 hpi	2	193	2.15	2.42	8977	191	1
13	Cuc 2 hpi	2	74.6	2.12	2.11	1339.5	28.5	4.1
14	Cuc 4 hpi	2	93.4	2.16	1.32	2660.2	56.6	4.1
15	Cuc 8 hpi	2	183.5	2.18	2.32	7849	167	3.4
16	Cuc 16 hpi	2	128.3	2.16	2.2	3647.2	77.6	1.1
17	Tom 2 hpi	3	109.4	2.06	1.75	4309.9	91.7	3.6
18	Tom 4 hpi	3	238.9	2.11	1.99	14476	308	3
19	Tom 8 hpi	3	323.1	2.09	2.05	8883	189	2.5
20	Tom 16 hpi	3	322.4	2.08	2.06	6705	149	1.9
21	Cuc 2 hpi	3	95.9	2.13	1.96	2401.7	51.1	3.5
22	Cuc 4 hpi	3	181.6	2.19	2.54	3290	70	3.2
23	Cuc 8 hpi	3	159.3	2.16	1.95	8178	174	2.5
24	Cuc 16 hpi	3	316.1	2.14	2.04	3332.3	70.9	1.5
25	Tom 2 hpi	4	114.1	2.04	1.45	3802.3	80.9	3.6
26	Tom 4 hpi	4	241.3	2.11	2.02	8554	182	3
27	Tom 8 hpi	4	347.3	2.07	1.96	1095.1	23.3	2.6
28	Tom 16 hpi	4	347.3	2.07	1.96	16074	342	1.8
29	Cuc 2 hpi	4	165.8	2.14	1.9	6016	128	3.7
30	Cuc 4 hpi	4	161.6	2.15	2.1	5217	111	3.4
31	Cuc 8 hpi	4	162	2.14	2.18	8178	174	1
32	Cuc 16 hpi	4	388.4	2.1	2.11	18236	388	1.4
33	GC 2 hpi	1	165.9	2.22	2.02	3391.2	94.2	5.3
34	GC 2 hpi	2	160.9	2.2	2.24	6786.8	144.4	5.4
35	GC 2 hpi	3	136	2.21	2.21	1809.5	38.5	5.3
36	GC 2 hpi	4	144.9	2.22	2.49	7567	161	4.8

germinating cyst samples (GC) which have an average RIN of 5.2 (Table 3.1). The RIN of RNA samples from zoospores that had been incubated in TE or CE were less that germinating cysts, there are several possible reasons for this. Either contamination with degraded RNAs from the plant extract, the extract itself causing

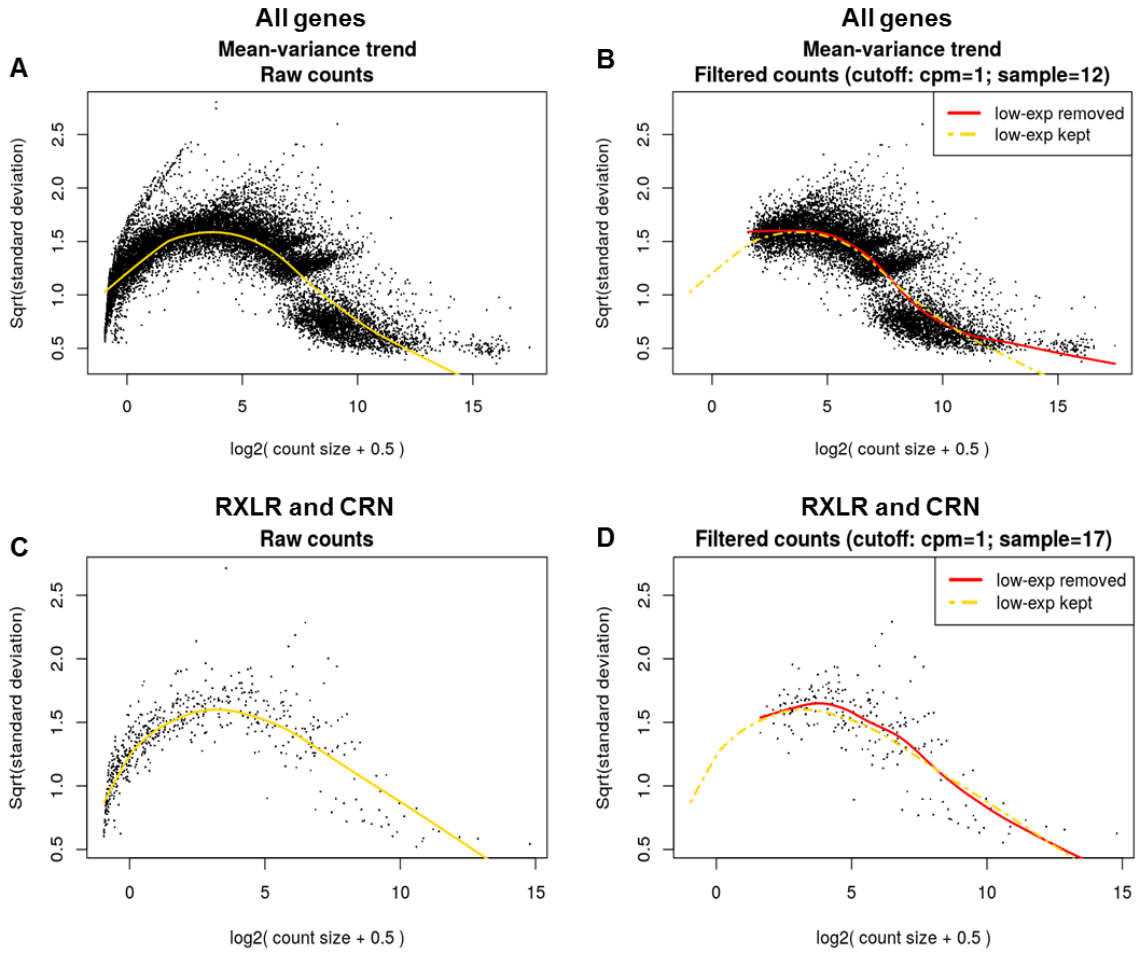


Figure 3.4: Mean variance trend, expression level of all expressed transcripts, and filtered transcripts.

The number of reads for each gene are assumed to follow a negative binomial distribution, with relation to variance plotted against the log2(number of read counts). The Mean varienc of the reads of gene for both the “all gene” database (A, B) and the effector (RXLR and CRN) database (C, D) are shown here raw (A, C) and post filtering (B, D). Reads were retained of an expressed transcript had ≥ 12 samples for “all genes “ and 17 for (RXLR and CRN) ≥ 1 (Counts per million) CPM expression.

some degradation during extraction (even though most has been removed), or a shift in the size of mRNAs when comparing GC to extracts. However, it is notable that all 16 hpi samples had a very low RIN and gel electrophoresis shows that very few high nucleotide length RNAs are present in the 16 hpi samples (Supplementary Table 1, Figure 3.5). These samples were therefore not taken forward for RNA-SEQ. It is difficult to explain the reason for this low integrity RNA from 16 hpi, perhaps due to

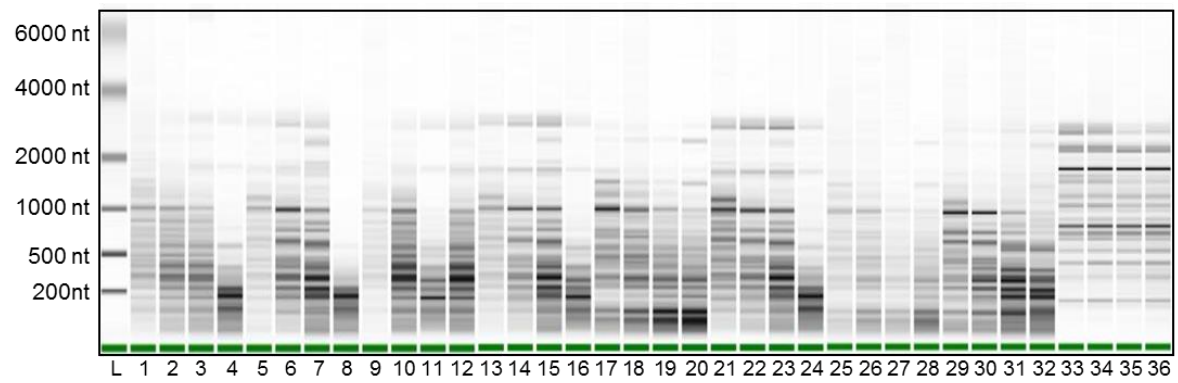


Figure 3.5: RNA gel electrophoresis.

Carried out by Novogene prior to RNA sequencing. All RNA samples numbered 1 – 36 are shown here (details of numbering shown in Table 3.1). 1 - 8 are in sequence Tomato extract (TE) 2 hpi, 4 hpi, 8hpi and 16hpi, followed by Cucumber extract (CE) 2 hpi, 4 hpi, 8hpi and 16 hpi. Thus 9 – 16 are the same sequence of samples and represent the second biological replicates, 17-24 the third biological replicates and 25 – 32 the fourth. Samples 33 – 36 are the four biological replicates of GC at 2 hpi.

decreased lysis of *P. capsici* tissue as by this time, there was high mycelial growth, meaning the tissue was more difficult to lyse.

3.2.3 Host extract induced differential gene expression

RNA-SEQ was conducted on samples from both host extracts (TE and CE) at 3 time points (2hpi, 4hpi and 8hpi) as well as germinating cysts in water at 2hpi (GC) used as a control (**Error! Reference source not found.**). Following data pre-processing and low expression transcript filtering, 9,974 genes were found to be expressed out of 19,729 originally annotated

by Salmon-quant (Galaxy Version 0.14.1.2) (Patro et al., 2017) (Figure 3.4). For this study, the interest lay, initially in the comparison between the two host plant extracts

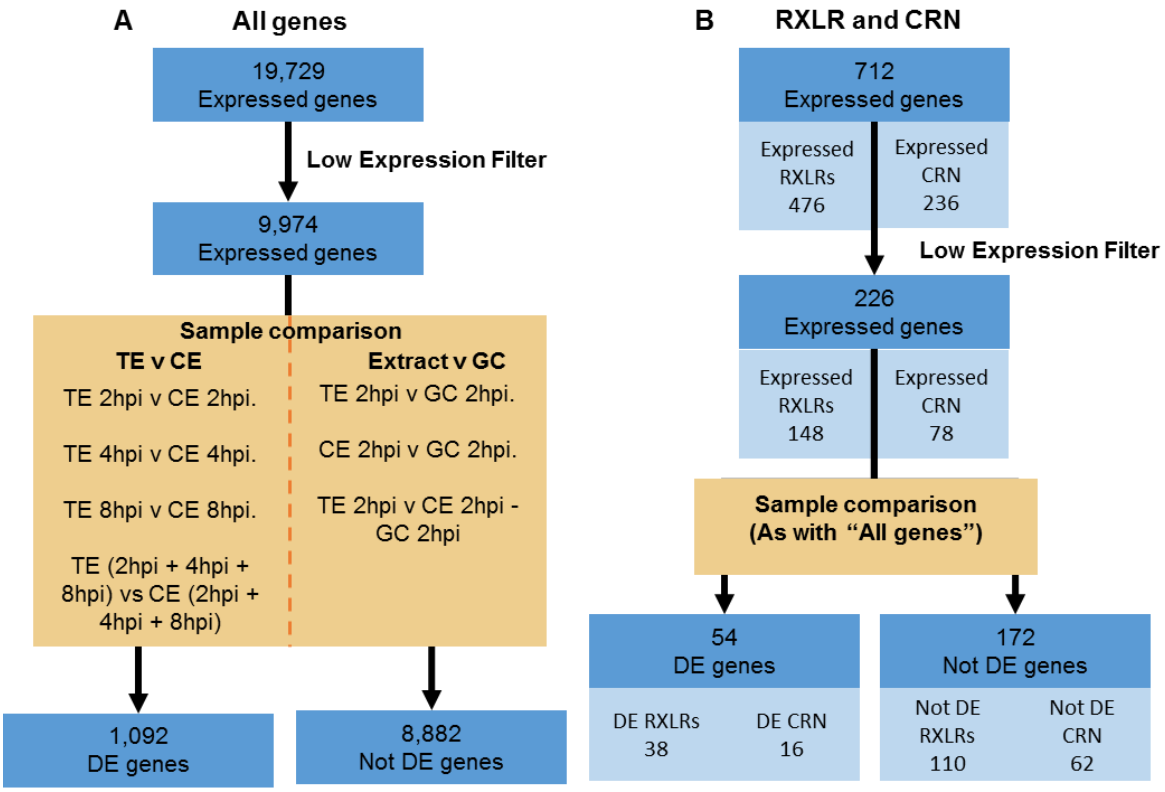


Figure 3.6: Number of gene expressed, before filtering, after filtering and after differential expression analysis.

Two databases of genes were used, all genes as annotated in the *P. capsici* LT1534 version 11 and an effector library containing RXLR and CRN genes. The number of genes that are found to be expressed in the RNA sequencing for both databases is shown, the number of each effector type is also shown. The number of genes following filtering of low expression gens is shown. Seven comparisons were then made of the purpose of differential expression analysis. The comparisons are shown for the All genes database but are the same for the Effector database, the comparisons have been split into two groups those between tomato extract and cucumber extract (TE v CE) and those between the host extracts and germinating cysts (Extract v GC). From these seven comparisons the number of differentially expressed (DE) genes and non-differentially expressed genes is shown.

(TE versus CE), and secondly in the ability of either host extract to elicit a response when compared to a water control (Extract v GC). The 7 comparison groups were split into two sets for results analysis, TE v CE and Extract v GC. For all comparisons, of the 9,974 expressed genes, 1,092 (10.9%) were found to be differentially expressed (DE) (Figure 3.6).

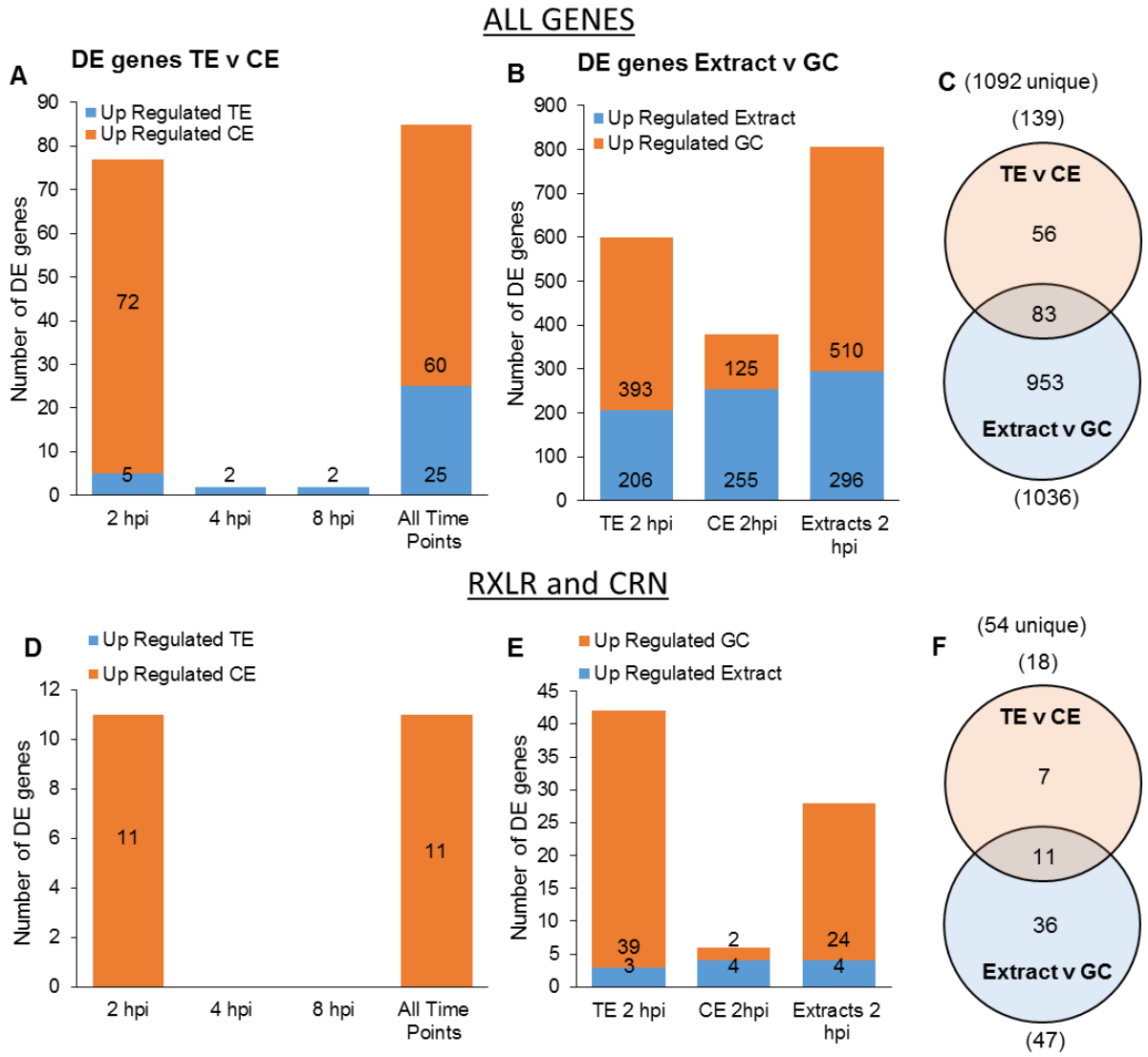


Figure 3.7: Number of differentially expressed genes in each comparison group.

For all genes as annotated in the P. capsici LT1534 version 11 (A-C), and the effector library containing RXLR and CRN genes (D-F) the number of differentially expressed (DE) genes for all seven comparisons are shown. The number of DE genes for the all gene database and the effector database are shown for the four tomato extract (TE) v cucumber extract (CE) (A, D). For these DE genes those up-regulated in TE are shown in blue on the lower portion of the graph and those genes up-regulated in CE are shown in orange the lower portion of the graph. In addition, the number of DE genes for the all gene database and the effector database are shown for the three host extract v germinating cysts (GC) (B, E). For these DE genes those up-regulated in Extract are shown in blue on the lower portion of the graph and those genes up-regulated in GC are shown in orange the lower portion of the graph. Note across some of the time points DE genes may appear in multiple comparison, Venn diagrams of the overlap of the DE genes from the four TE v CE and the three extracts v GC are shown for the all gene database (C) and effector database (F).

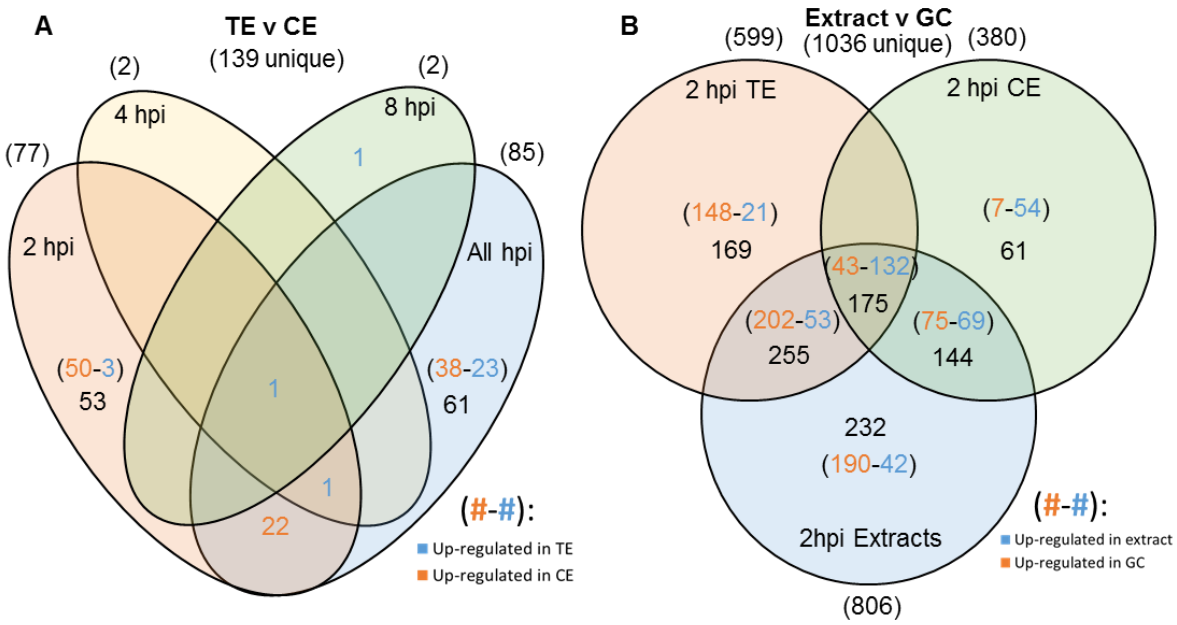


Figure 3.8: Venn diagrams showing the overlap of each comparison group in both the TE v CE set and the Extract v GC set.

The overlap all 139 of the unique DE genes from the tomato extract comparison with cucumber extract (TE v CE) is shown (A). The number of genes total is shown in brackets above the section. For the four TE v CE comparison (2 hpi, 4 hpi, 8 hpi, All hpi) the total number of gene in each overlap is shown, those gene that are up-regulate in TE are shown in blue, and those gene up-regulated in CE are shown in orange. The overlap all 1036 of the unique DE genes from the two host extract vs germinating cysts (Extract v GC) is shown (B). The number of genes total is shown in brackets above the section. For the three Extract v GC comparison (TE 2 hpi, CE 2 hpi, and both extracts pooled 2 hpi) the total number of gene in each overlap is shown, those gene that are up-regulate in Extract are shown in blue, and those gene up-regulated in GC are shown in orange.

3.2.4 Host extract induced differential gene expression: Extracts compared with germinating cysts

Of these 1,092 DE genes, the majority (1,036, 94.9%) are differentially expressed when comparing between transcript levels in host extracts and GC (Figure 3.7). The results show that the greatest difference in gene expression is shown when comparing GC to TE at 2hpi, with 599 differentially expressed genes (DEGs), in contrast to the

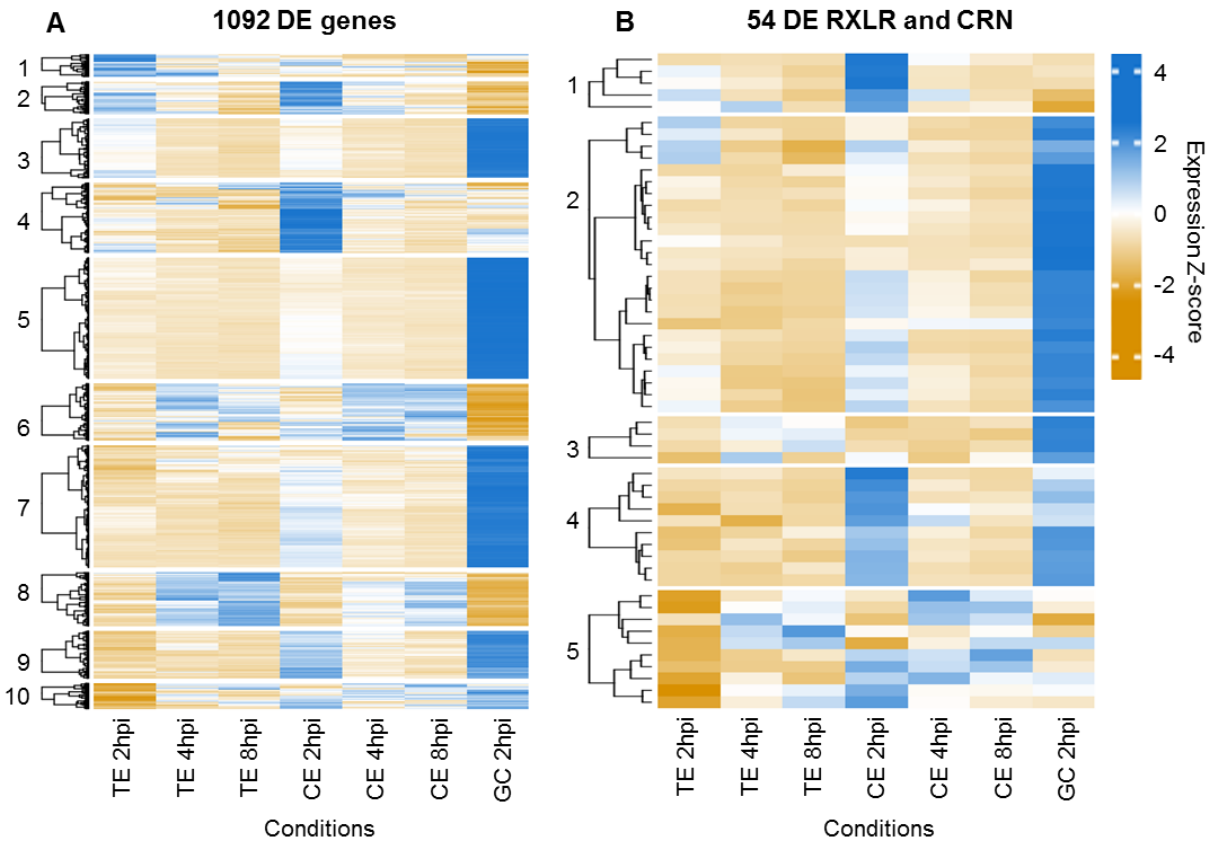


Figure 3.9: Hierarchical cluster analysis of all differentially expressed genes.

The DE genes from all seven comparisons for all genes from the database Two databases of genes were used, all genes as annotated in the P. capsici LT1534 version 11 (REF) and an effector library containing RXLR and CRN genes. Pearson's correlation coefficient was used to for Hierarchical clustering of the genes and for all genes 10 clusters are shown on the right, and for the RXLR and CRN genes 5 cluster have been calculated. The expression is shown by Z-score.

380 DEGs when comparing GC to CE at 2 hpi (Figure 3.7). It is notable that for those uniquely DE genes between GC and TE at 2 hpi, the majority are induced in GC, 350 genes contrasted with 74 genes up-regulated in TE 2 hpi (Figure 3.8). Additionally, those genes that are DE and shared between both host extract comparison-groups (TE 2hpi compared with GC, and CE 2hpi compared with GC) show the opposite pattern, the majority of the genes are induced in both the host extracts, with 132 DEGs up-regulated by both extracts, and only 43 genes up-regulated in GC. Interestingly, and differing from TE 2 hpi, unique DE genes in CE 2 hpi show a similar but less

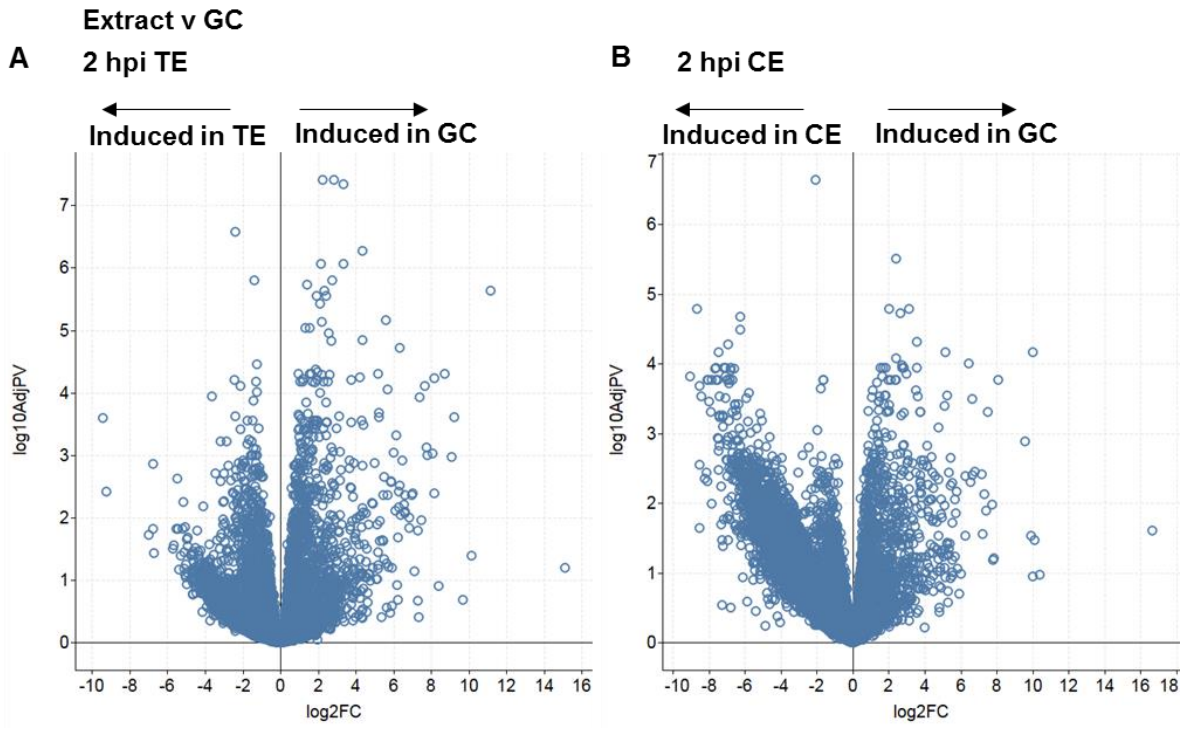


Figure 3.10: volcano plot show DE of genes at 2hpi TE and CE both compared to 2hpi GC.

The log2 fold change (FC), and the Log10 p value (adjusted with Bonferroni correction) for each gene when comparing between two conditions has been plotted. The two comparisons were Tomato extract (TE) at 2 hours post inoculation (hpi) compared with germinating cysts at 2 hpi (GC) (A), at cucumber extracts at 2 hpi (CE) compared with GC (B).

pronounced pattern, with 123 genes being induced in CE to the 82 up-regulated in GC. Both host extract expression data combined (2hpi Extract) and compared to GC gave an additional 232 DE genes which showed a similar pattern to those genes unique to TE 2 hpi: indeed, of the 232 genes, 190 were up-regulated in GC and only 42 were up-regulated in 2hpi Extract (Figure 3.8).

3.2.5 Host extract induced differential gene expression: Host to Host comparison

DE of genes was also found when gene expression in TE was compared to CE. Host-host comparisons were made at 2 hpi, 4 hpi, 8 hpi, and at every time point pooled (All-

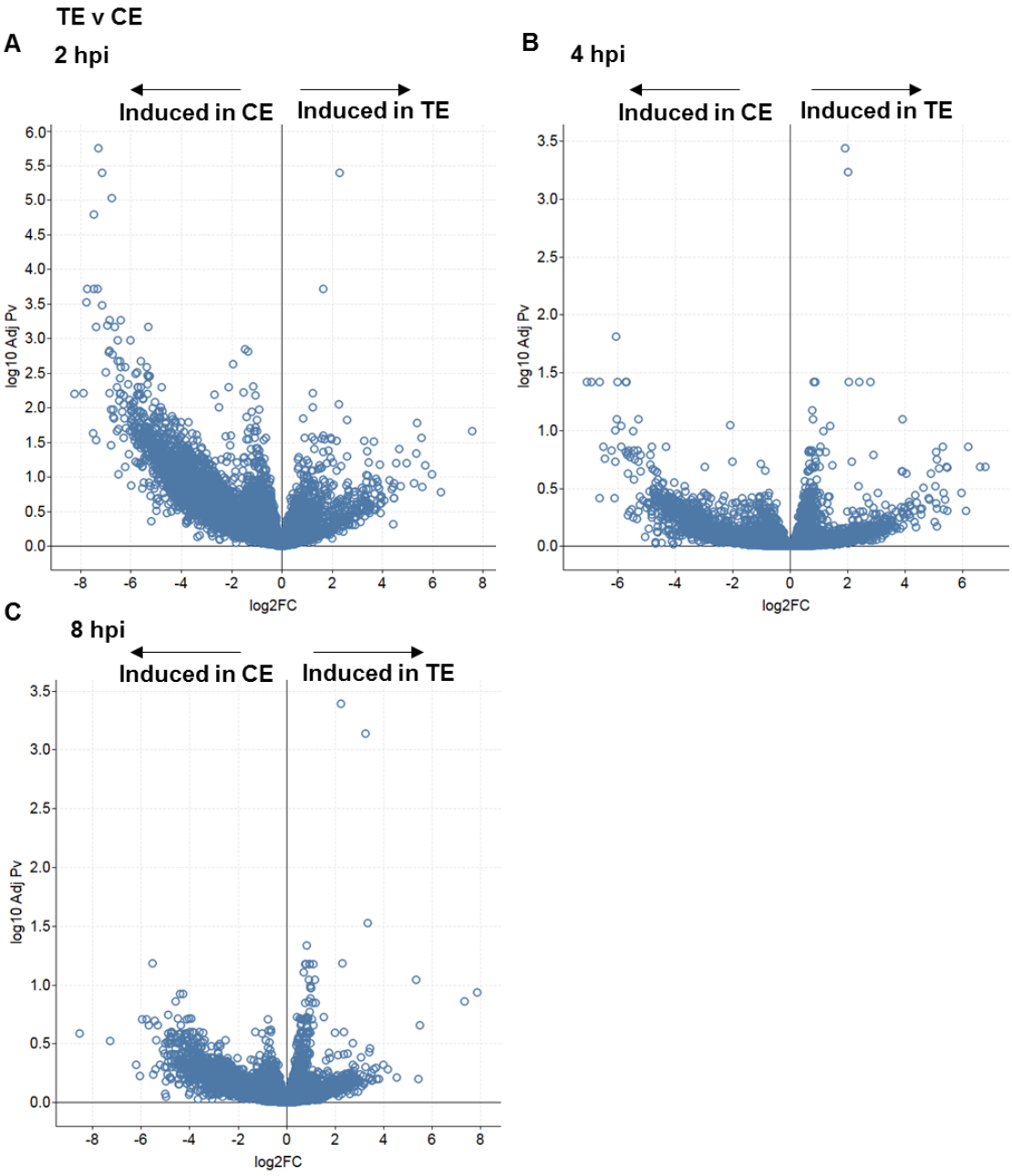


Figure 3.11: volcano plot show DE of genes at 2hpi, 4hpi and 8hpi comparing TE with CE.

The log2 fold change (FC), and the Log10 p value (adjusted with Bonferroni correction) for each gene when comparing between two conditions has been plotted. Here three comparisons were made: Tomato extract (TE) at 2 hours post inoculation (hpi) compared cucumber extracts at 2 hpi (CE), TE 4 hpi compared with CE 4hpi, and TE 8 hpi compared with CE 8 hpi.

hpi). Of the 9,974 genes expressed only 139 (1.4%) were differentially expressed between the two hosts (Figure 3.7). However, it was striking that of the 139 DE genes

110 (79.1%) of them had higher expression in CE; All 110 genes were found to have differentially higher expression in CE at 2 hpi, or in All-hpi, (50 found uniquely at 2 hpi, 22 found at both 2 hpi and All-hpi and 38 only found when the data for all time points were pooled). Of the remaining 29 genes that were DE, and had higher expression in TE, the majority (23) were only DE at All-hpi. Interestingly, there was a larger temporal spread in *P. capsici* genes with higher expression in TE, a spread that did not occur in CE where all genes were DE at 2 hpi or in pooled data (Figure 3.8). In summary TE did not show a great differential increase in transcript levels, either when compared to CE or when compared to GC, contrastingly CE and GC show a much larger number of transcripts with increased expression levels.

A similar pattern can be found if DE is separated by hierarchical cluster analysis. The 1,092 DE gene were separated into 10 Clusters. The induction of genes in GC and CE can be easily spotted. Clusters 3, 5 and 7 where there is marked induction in GC, in Clusters 2 and 4 where genes induced in CE can be found in and in Clusters 9 and 10 where those genes are collected that are both induced in CE and GC. These clusters represent 49.5% (Clusters 3, 5 and 7), 16.8% (Clusters 2 and 4) and 12% (Clusters 9 and 10) of the total DE gene suite. This leaves only Cluster 1, 6 and 8 where there is increased gene expression in TE. Cluster 1, the smallest cluster (40 genes) are those gene induced in TE, particularly at 2 hpi, and Clusters 6 and 8 (18.1% of total DE genes) are genes showing a shared induction in TE and CE and down-regulation in GC. Interestingly, in these clusters, shared induction in TE and CE for the majority happens in 4 hpi and 8 hpi (Figure 3.9). A further breakdown of comparison group clusters at different stages of inoculation can be found in table 2 (Supplementary Table 2)

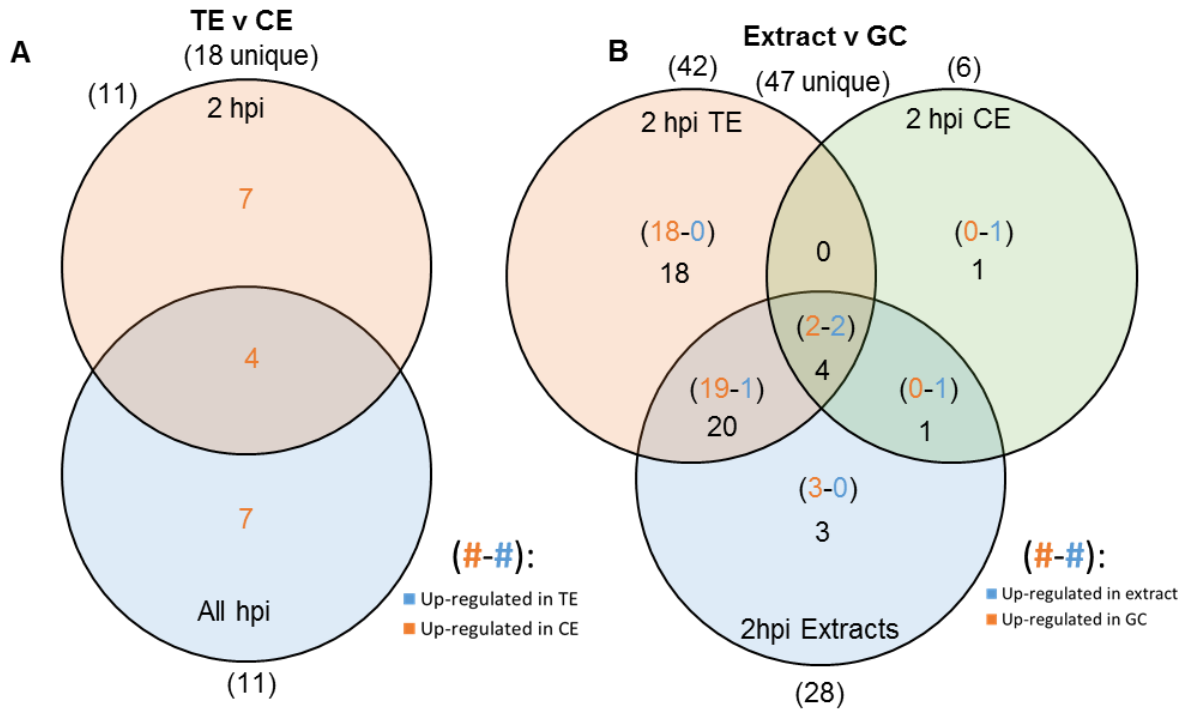


Figure 3.12: Venn diagram showing overlap in DE RXLR and CRN genes in both the TE v CE set and the Extract v GC set.

The overlap all 18 of the unique DE genes from the tomato extract comparison with cucumber extract (TE v CE) is shown (A). The number of genes total is shown in brackets above the section. For the four TE v CE comparison (2 hours post inoculation (hpi), 4 hpi, 8 hpi, All hpi) only two had DE expressed genes (2hpi and All hpi) the total number of gene in each overlap is shown, those gene that are up-regulate in TE are shown in blue, and those gene up-regulated in CE are shown in orange. The overlap all 47 of the unique DE genes from the two host extract vs germinating cysts (Extract v GC) is shown (B). The number of genes total is shown in brackets above the section. For the three Extract v GC comparison (TE 2 hpi, CE 2 hpi, and both extracts pooled 2 hpi) the total number of gene in each overlap is shown, those gene that are up-regulate in Extract are shown in blue, and those gene up-regulated in GC are shown in orange.

Upon inoculation of CE, there is a marked increase in the expression of genes at 2 hpi that is not seen in TE. This shift is lost as inoculation continues and there is little significant differential expression of genes at 4 hpi and 8 hpi (Figure 3.11). However, host extracts, TE and CE, seem to be able to cause a large number of changes in the expression profile of *P. capsici* as can be seen when comparing to GC at 2 hpi. However, there is a host difference in the type of response. The difference between

TE and GC is characterised by a large induction of genes in GC and limited induction of genes in TE. In contrast to this, the response of *P. capsici* zoospores to CE when compared to GC shows a greater induction of genes in CE than in GC. However, the difference between extract and GC for CE versus GC in terms of induction in one condition compared to the other is not as marked as it is in TE versus GC (Figure 3.10)

3.2.6 Host extract induced differential RXLR and CRN expression

In addition to the 19,729 total annotated genes, transcripts from another set of genes were also annotated and quantified using the salmon quant tool (Galaxy Version 0.14.1.2) (Patro et al., 2017). This set included both known and putative effector protein genes, RXLRs and CRNs. All 712 were found to be expressed at some level, however, after filtering low abundancy transcripts 226 RXLRs and CRNs were found to be expressed (Figure 3.4), of these 148 were RXLRs and 78 were CRNs. Using the same 7 comparison groups and with all other genes again separated into two sets for analysis, TE v CE and Extracts v GC, 52 of these 226 (23%) were found to be differentially expressed. Of the 52 DE effectors 38 were RXLRs and 16 were CRNs (Figure 3.6).

In terms of expression pattern, the RXLRS and CRNs display similar patterns to the other genes. The majority of genes were found DE when comparing the two host extracts to GC. 47 genes were DE comparing Extract to GC this is in contrast with the 18 genes DE when comparing TE to CE, 11 genes are shared in these two sets of DE effectors. The majority of these 47 DE effectors are induced in GC, more so when the GC are compared to TE, few genes are induced in GC when comparing to CE. This shows again the up-regulation events that occur both in GC and CE (Figure 3.7, Figure 3.12). Interestingly, the effectors that are induced in the extract are either found

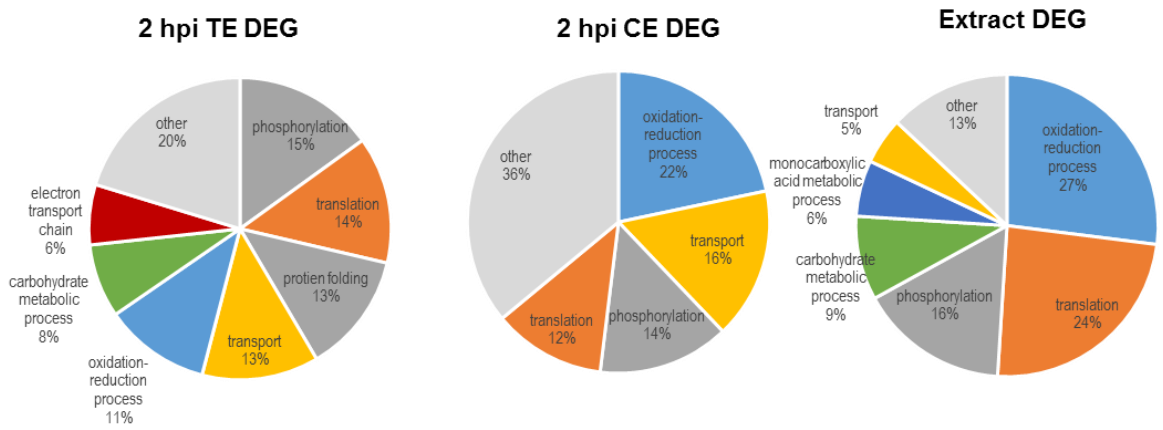


Figure 3.13: the level of each GO term in genes DE when comparing 2hpi extracts to GC, in TE, CE and in Both combined compared to GC, both incuced in extract and incued in GC.

The GO (gene ontology) terms of all the DE genes from the comparison of host extracts at 2 hours post inoculation (hpi) and germinating cysts at 2 hpi were obtained. The GO term of the differentially expressed genes (DEG) for the category of biological processes have been separated down by comparison TE 2hpi v GC, CE 2hpi v GC and both extracts pooled v GC. Thus the percentage of the top biological process GO terms form each comparison is shown in the pie charts.

uniquely induced in 2hpi CE or shared between 2 hpi CE and TE. CE seems able to induce gene expression in *P. capsici* zoospore whereas DE in TE is characterised by down regulation events. This is shown when comparing TE to CE for effector expression. 11 RXLRs or CRNs are DE when comparing host extracts, all are induced in CE and the most are found during the 2 hpi time point (Figure 3.7, Figure 3.12).

Hierarchical cluster analysis of the 54 DE RXLRs and CRN into 5 clusters revealed their expression patterns. The largest cluster (Cluster 2) are effector genes induced in GC, additionally, Cluster 3 also displays a similar expression pattern. Cluster 4 and Cluster 1 are also notable having genes up-regulated in CE 2 hpi and GC and those genes up-regulated just in CE 2 hpi. In total 87.0% of genes are found in clusters 1-4 and therefore either induced in GC, CE 2 hpi or both. The final cluster, cluster 5 are genes that are notably turned off in 2 hpi TE and show variable induction at other time points, again showing the TE is characterised by down-regulation events (Figure 3.9).

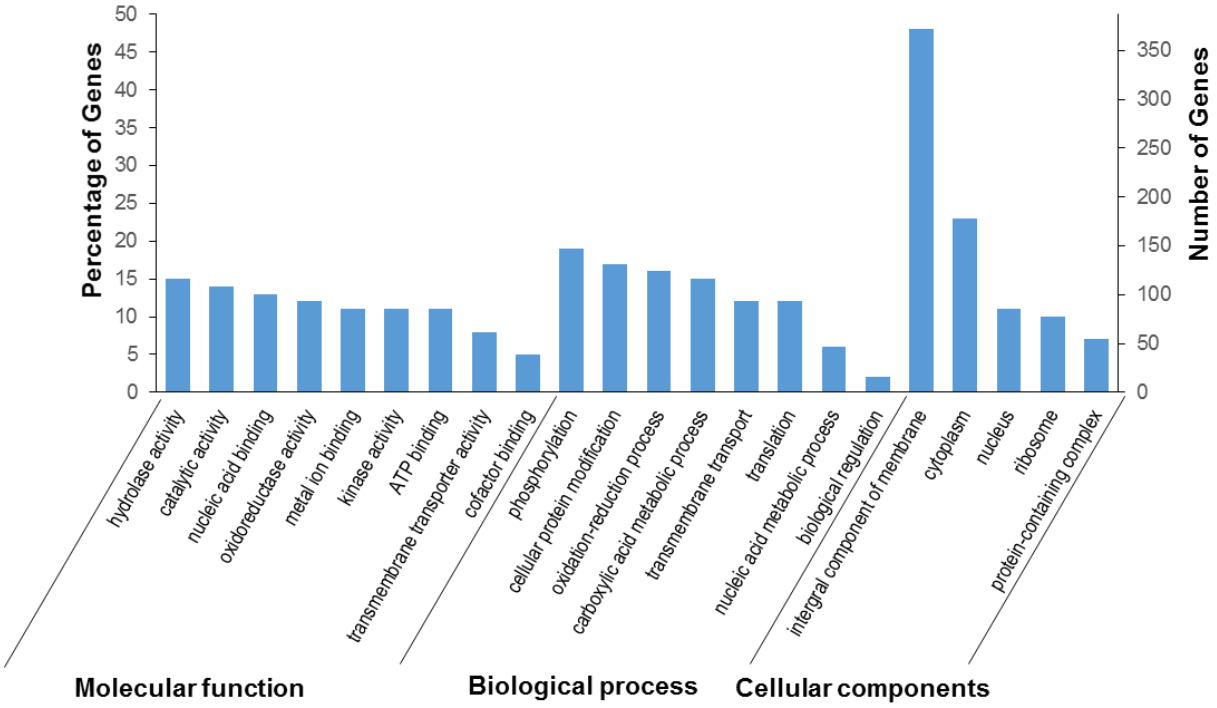


Figure 3.14: the level of each GO term in genes when comparing 2hpi extracts to GC.

The GO (gene ontology) terms of all the DE genes from the comparison of host extracts at 2 hours post inoculation (hpi) and germinating cysts at 2 hpi were obtained. The top 9 GO terms for all the genes from the Molecular function category and both the percentage of DE genes and the number of DE genes that have be annotated with that GO term is shown. In addition the top 8 GO terms from the biological function category and the top 5 genes from the cellular component category and there respective percentages and numbers of DE genes are also shown.

3.2.7 DE Gene characterisation using GO term annotation

The role of the putative effector gene is assumed to be as a normal effector, i.e. to suppress the plant immune system or allow extraction of nutrients, and either being secreted into the plant cell or apoplastic space. However, the function of the other 1092 DE genes are unknown. GO term analysis of DEGs was conducted to glean gene function and identify the mechanisms that could contribute to dynamic host adaption. First, the 1036 genes that were DE when comparing extract to GC were GO annotated. GO terms are broken up into three main classes, molecular function,

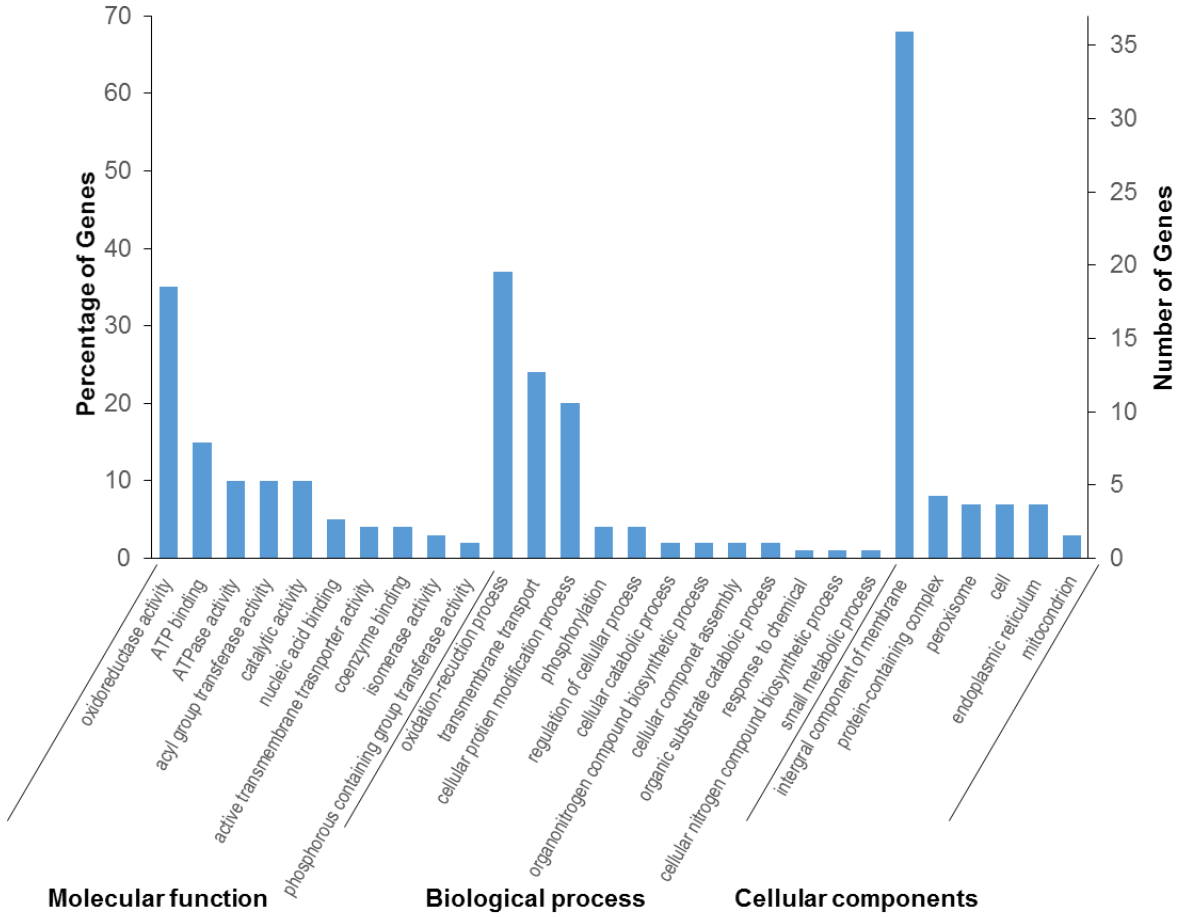


Figure 3.15: the level of each GO term in genes DE when comparing TE to CE.

The GO (gene ontology) terms of all the DE genes from the comparison of tomato extracts (TE) at 2 hours post inoculation (hpi) and cucumber extract (CE) at 2 hpi were obtained. The top 10 GO terms for all the genes from the Molecular function category and both the percentage of DE genes and the number of DE genes that have be annotated with that GO term is shown. In addition the top 12 GO terms from the biological function category and the top 6 genes from the cellular component category and there respective percentages and numbers of DE genes are also shown.

biological process, and cellular component. Of the GO terms that are classed as molecular function, amongst the most enriched were hydrolase activity (GO:0051336) catalytic activity (GO:0003824), nucleic acid binding (GO:0003676) and oxidoreductase activity (GO:0016491). In terms of biological processes GO terms that were enriched, the three most common were phosphorylation (GO:0016310), cellular protein modification (GO:0036211) and oxidation-reduction processes (GO:0055114). Interestingly, when examining the GO terms for the cellular component, by far the

largest group of genes are annotated as integral components of the membrane (Figure 3.14). Breaking down the distribution of these GO terms into the three comparison groups included in the extract v GC set, and just examining those up-regulated in Extracts (TE, CE, and Shared) reveals something more. In those genes that are uniquely up-regulated in CE, or DE up-regulated in both extract set, that largest set of genes have been annotated with GO terms for oxidation-reduction processes. Additionally, the second most common GO term in CE is transmembrane transport (GO:0055085) and is not found to be enriched in the shared genes or TE. Contrastingly, the largest groups in TE DE genes are phosphorylation, translation (GO:0006412), protein folding (GO:0006457) and transport (GO:0006810). Phosphorylation and translation are not unique to TE, however, and they are the 3rd and 4th most common in gene up-regulated in CE and 2nd and 3rd most common in shared genes (Figure 3.13).

When comparing TE to CE and looking at those genes just DE, the picture is clearer. The majority of these genes are up-regulated in CE. The common five GO terms in this set are oxidoreductase activity, oxidation-reduction processes, transmembrane transport, cellular modification processes and integral component of the membrane. A similar list of potential roles to those identified when examining DE gene from the comparison of CE with GC (Figure 3.15).

3.2.8 Expression of known genes and marker genes

The expression profile gained via RNA-SEQ of several marker genes and genes of interest was examined. This may provide some information mechanistically on what is occurring inside *P. capsici* and examining genes with known expression pattern can enable validation of the sequencing. The expression of some common *P. capsici*

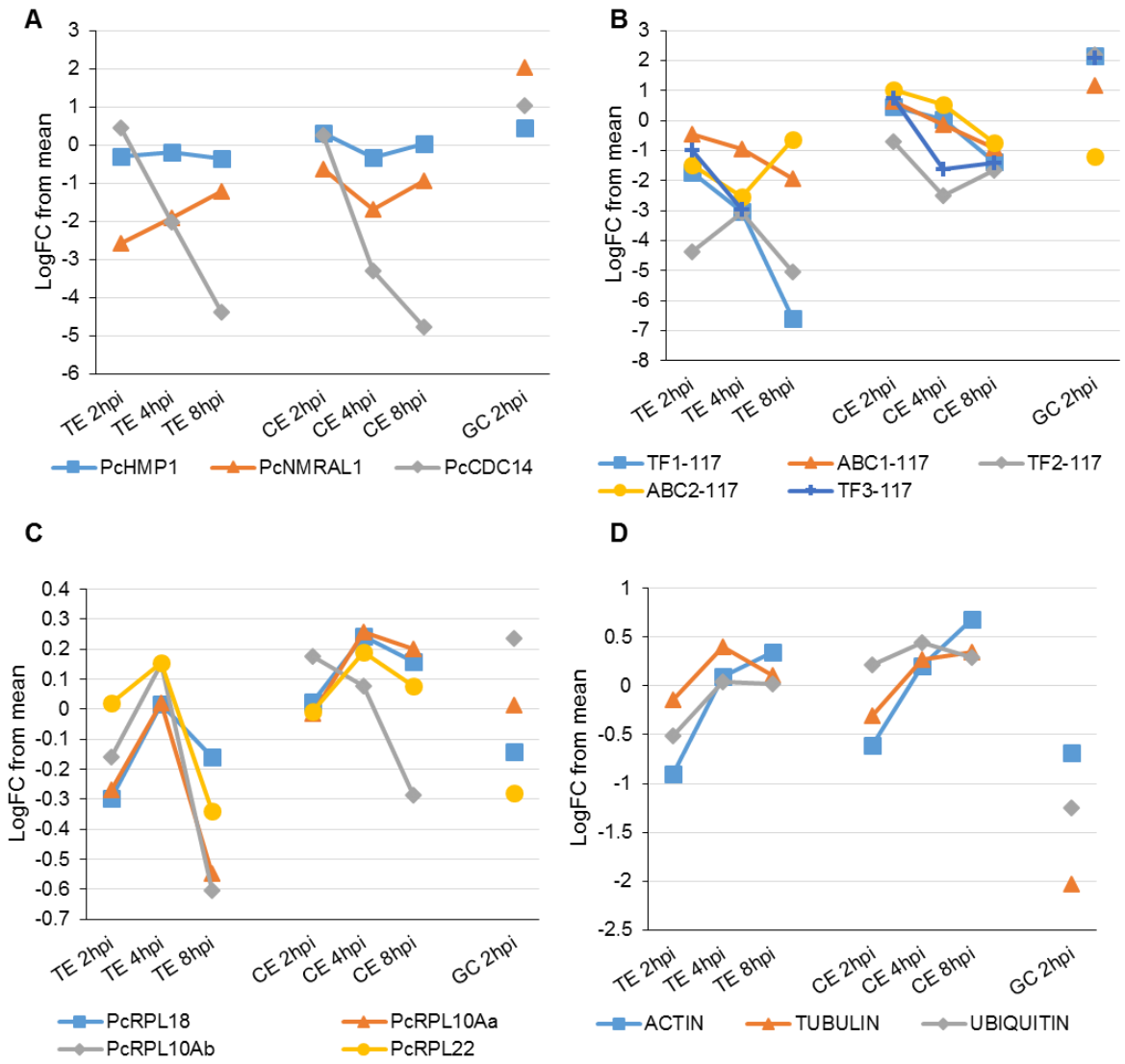


Figure 3.16: expression pattern as shown by RNA seq of key genes.

For several key genes, consisting of three marker genes PcHMP1, and PcCDC14, and PcNMRAL1 (A), 5 gene thought to be differentially regulated by host TF1-117, ABC1-117, TF2-117, ABC2-117 and TF3-117 (B). Four large ribosomal protein subunits PcRPL18, PcRPL10Aa, PcRPL10Ab, and PcRPL22 (C), and three so called housekeeping genes ACTIN, TUBULIN and UBIQUITIN (D). Relative Fold change for each gene from the mean expression of that gene was calculated. Expression in the tomato extract (TE) and cucumber extract (CE) at the three time points 2 hours post inoculation (hpi), 4 hpi, and 8 hpi, and in germinating cyst (GC) 2 hpi.

marker genes were analysed. PcHMP1, a marker of biotrophy associated with haustoria formation is uniformly expressed in all time points and extracts, as well as GC. PcNMRA1 (Phyca11_505845) which is switched off during necrotrophy and as is known to mediate the switch from biotrophy to necrotrophy (Pham et al., 2018) is also

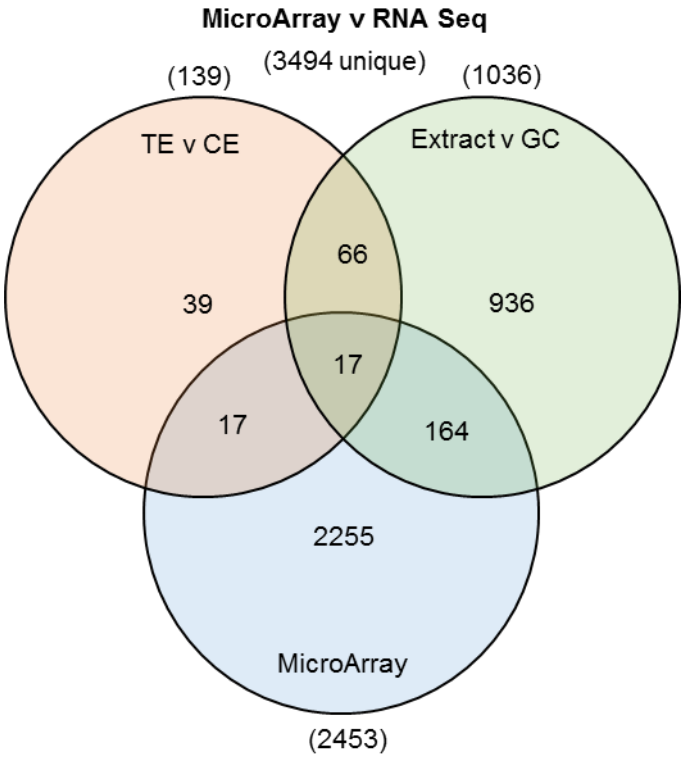


Figure 3.17: Venn diagram showing overlap in DE gene identified from RNA seq of zoospore inculcation experiments and 16hpi microarray of leaf infection assay of four host plants.

Differentially expressed (DE) genes from the two main groups of comparisons in the RNA seq experiments, Tomato extract v cucumber extract (TE v CE) and Host extracts versus germinating cysts (Extract v GC) were collected. The overlap of all these DE genes with those from the microarray is shown, the total number of unique total number DE genes from all experiments is shown at the top and the number of DE genes in each category is shown in brackets beside the Venn diagram.

fairly uniformly distributed, it is mildly induced in GC. Expression of PcCDC14 (Phyca11_ 510939), a marker of sporulation, was also examined, this showed rapid down regulation as the time progressed in both extracts (Figure 3.16).

The expression profiles of genes frequently used as housekeeping genes were examined, these genes ought to have stable expression. However, their expression is fairly different in the different samples, all housekeeping genes tested actin (Phyca11_ 132086), tubulin (Phyca11_ 576734) and ubiquitin (Phyca11_510705), showed increased expression as the experiment progressed, as well as low expression in GC

samples. Worryingly for most housekeeping genes, a high level of variability was seen. For example, tubulin showed a range of 0.5 fold change above the mean to a twofold decrease below mean expression. In addition to these genes, other genes with presumably high stable expression were analysed, four large ribosomal protein (RPL) subunit gene expression was also examined (PcRPL18 – Phyca11_ 9769, PcRPL10Aa – Phyca11_15914, PcRPL10Ab – Phyca11_ 39189, and PcRPL22 Phyca11_ 19566). These RPLs had a much more stable expression pattern, with the largest range displayed in RPL10Ab (0.23 fold change above the mean to 0.60 fold change below the mean) and the smallest range in PcRPL22 (0.19 FC above the mean to 0.34 FC below the mean) (Figure 3.16).

Finally, the expression of some genes of interest that were used during the validation test were examined. All the genes have previously shown differential gene expression, specifically up-regulation in cucumber infection. Validation of the *in vitro* methodology has also shown these genes to be up-regulated in CE in a similar fashion. The RNA-SEQ data concurs with all of the other results, that there is an up-regulation of these genes in CE compared to TE. However, the RNA-SEQ data, as well as some of the previous validation experiment, also showed that GC had higher expression than most of these genes of interest. Having said that, none of these genes were significantly differentially expressed in comparison between the two host extracts (Figure 3.16).

In fact, comparing the genes that are differentially expressed in this RNA-SEQ experiment and those that appear differentially expressed according to microarray data of 4 hosts leaf infection assay at 16hpi, shows that there is little overlap. (Unpublished Huitema lab data). In total 198 genes are shared between the two experiments (Figure 3.17).

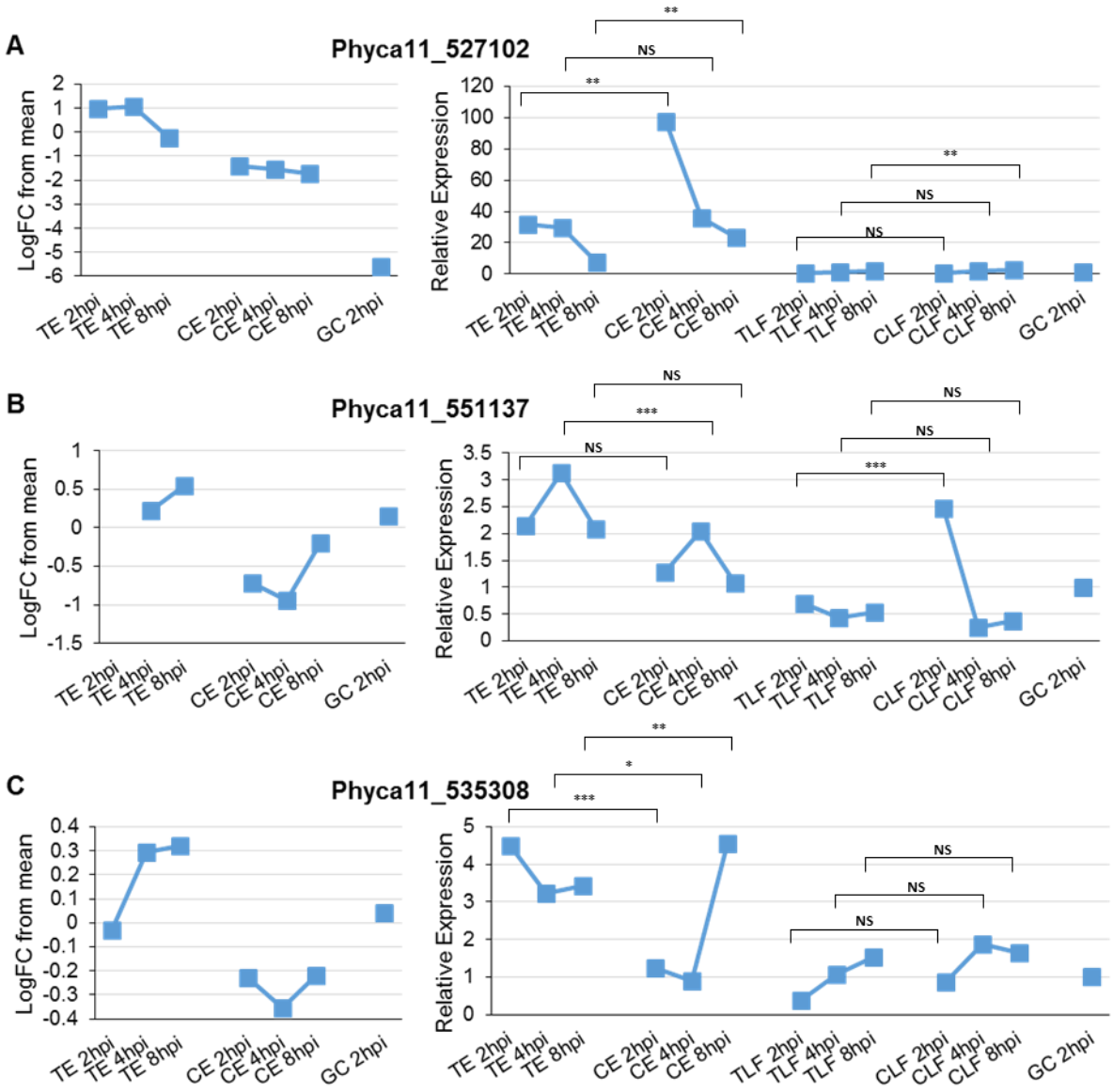


Figure 3.18: Validation experiments, comparison of expression pattern of genes shown by RNA seq and qRT-PCR (1).

Three genes that were shown to have differential expression (DE) in the RNA seq were probed for expression in Tomato extract (TE) 2 hours post inoculation (hpi), TE 4 hpi, TE 8 hpi, cucumber extract (CE) 2 hpi, CE 4 hpi, CE 8 hpi, tomato leaf infection (TLF) 2hpi, TLF 4 hpi, TLF 8 hpi, cucumber leaf infection (CLF) 2 hpi, CLF 4hpi, CLF 8hpi, and germinating cysts at 2 hpi. Relative expression calculated using the ΔΔCT method against the housing keeping gene RPL22, with the expression of GC 2hpi set to 1 is shown for three genes Phyca11_527102 (A), Phyca11_551137 (B), and Phyca11_535308 (C). The expression pattern, as shown By the Log fold change from the mean expression of that gene, from the RNA seq experiment is also shown for all these genes. Statistical significance is shown for comparison between TE and CE at 2hpi, 4hpi and 8hpi, and TLF and CLF at 2hpi, 4hpi and 8hpi. Two tailed unequal variance student t-test was conducted, significant bonferroni adjusted P values are shown. Adjusted P values of <0.05 = *, <0.01 = **, <0.001 = ***, >0.05 = non-significant (NS).

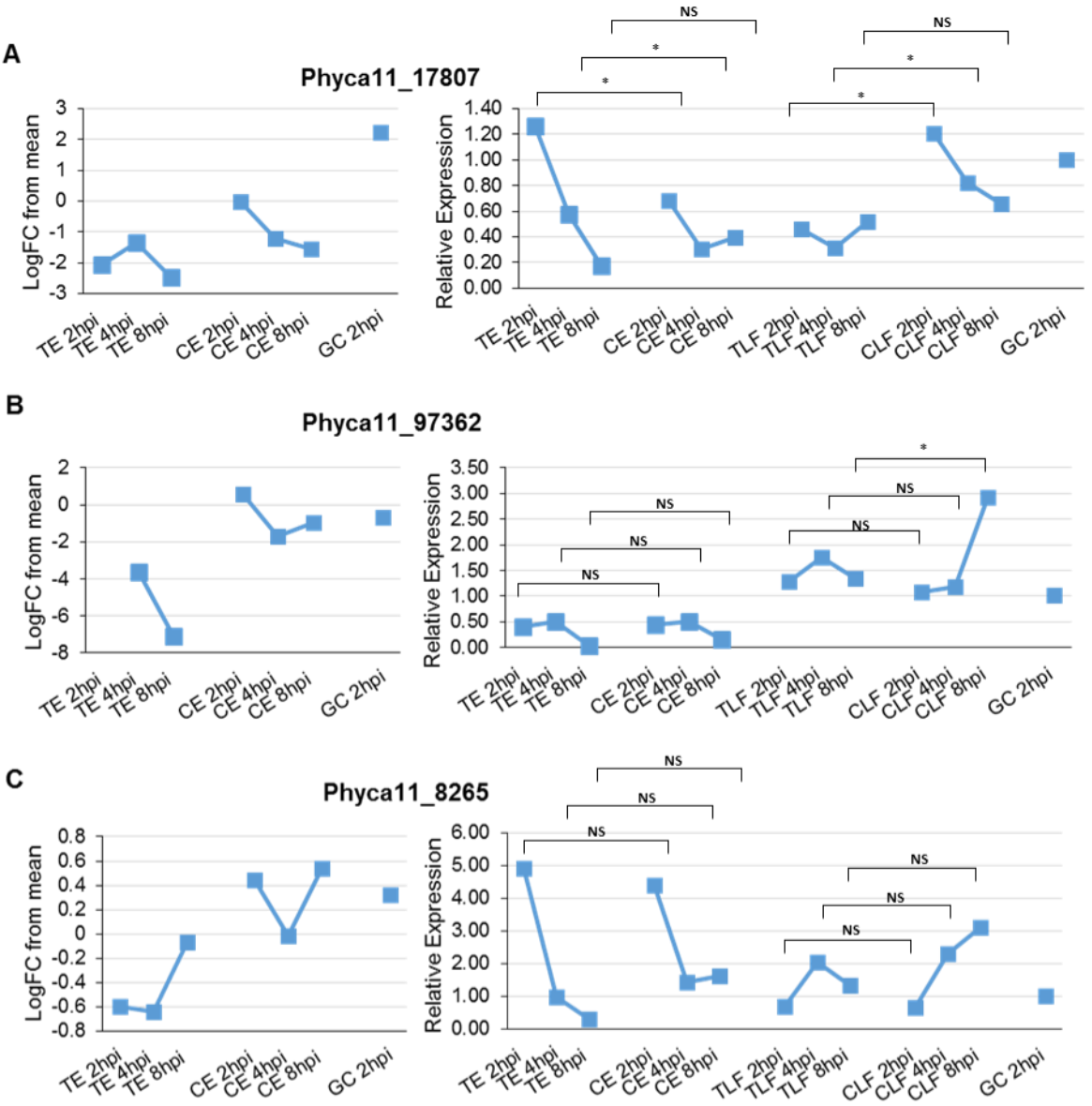


Figure 3.19: Validation experiments, comparison of expression pattern of genes shown by RNA seq and qRT-PCR (2).

Three genes that were shown to have differential expression (DE) in the RNA seq were probed for expression in Tomato extract (TE) 2 hours post inoculation (hpi), TE 4 hpi, TE 8 hpi, cucumber extract (CE) 2 hpi, CE 4 hpi, CE 8 hpi, tomato leaf infection (TLF) 2hpi, TLF 4 hpi, TLF 8 hpi, cucumber leaf infection (CLF) 2 hpi, CLF 4hpi, CLF 8hpi, and germinating cysts at 2 hpi. Relative expression calculated using the ΔΔCT method against the housing keeping gene RPL22, with the expression of GC 2hpi set to 1 is shown for three genes Phyca11_17807 (A), Phyca11_97362 (B), and Phyca11_8265 (C). The expression pattern, as shown By the Log fold change from the mean expression of that gene, from the RNA seq experiment is also shown for all these genes. Statistical significance is shown for comparison between TE and CE at 2hpi, 4hpi and 8hpi, and TLF and CLF at 2hpi, 4hpi and 8hpi. Two tailed unequal variance student t-test was conducted, significant bonferroni adjusted P values are shown. Adjusted P values of <0.05 = *, <0.01 = **, <0.001 = ***, >0.05 = non-significant (NS).

3.2.9 Validation

To validate the RNA-SEQ experiments, the expression pattern observed was attempted to be replicated with qRT-PCR. Six genes of interest were tested, three with higher expression in TE (Figure 3.18), and three with higher expression in CE (Figure 3.19), however only one replicate is present in each experiment, so conclusions that can be drawn from these experiments are limited. In addition, these experiments also contain leaf infection assay samples from the same time points as extract inoculation samples were taken. For two out of three genes induced in TE (Phyca11_551137 and Phyca11_535308) the differential host expression pattern is maintained, however the expression profile as the experiment progresses is different, i.e. they are induced/ down-regulated at different time points in the experiment. For the other gene induced in TE (Phyca11_527102) though the expression pattern is similar across time points for each extract, the host to host relative expression is different (Figure 3.18). For all three genes induced in TE, the expression in leaf infection assay is lower than that in extract. The only exception is the high level of expression on Phyca11_551137 at 2 hpi cucumber leaf infection assay (Figure 3.18). For the three genes that showed higher expression in CE in the RNA-SEQ, none of the host specific differential expression patterns were maintained. Of the three Phyca11_17807 shows the closes resemblance, although a high expression in TE in the qRT-PCR at 2 hpi limits the similarity. Interestingly though, for all of the genes, but especially Phyca11_17807 the host specific differential expression pattern observed in the RNA-SEQ is present in the leaf infection assay as shown by qRT-PCR (Figure 3.19).

3.3 Methodology

3.3.1 Broths-Extracts-Sandwiches.

Commercially available tomato plants (*Solanum Lycopersicum* cv. Moneymaker) and cucumber plants (*Cucumis sativus* cv. Venlo Pickling) susceptible to *P. capsici* were grown with a 16 hour light cycle and maintained at 22-25°C. 3 – 5-week old tomato and cucumber detached leaves were used for leaf infection assays and sandwich infection assays and leaf and stem material were used for broth and extract inoculation assays. *P. capsici* strain LT1534 was grown on V8 agar (10% v/v V8 vegetable juice, 1 g/L calcium carbonate, 300 mg/L β-sitosterol, 15% w/v Agar). For the purposes of infection and inoculation assays, zoospores were collected. *P. capsici* LT1534 was grown on V8 agar in the dark at 25°C sealed with parafilm. This was followed by 2 days of growth in the light at 22°C without parafilm to induce sporangia formation. Zoospores were collected by flooding the 150 mm plate with 30 mL of sterile distilled water at room temperature, then the mycelia growth was agitated with a sterile plate spreader and everything was transferred to a second *P. capsici* inoculated V8 agar plate. Continuing in this fashion, the 30 mL of water was used to flood and agitate four 150 mm plates recovered into a 50 mL falcon tube. The sporangia suspension was left for 30-45 minutes with the lid removed on a light-box to induce zoospore release. The suspension was then filtered through one layer of Miracloth to remove mycelia and any agar chunks in the suspension, however both full and empty sporangia are retained, this process yielded approximately 25 mL of ≥ 1 x 10^6 spores mL^{-1} zoospore suspension. The zoospore suspension was then diluted to 1 x 10^6 spores mL^{-1} (**Error! eference source not found.**). For the purposes of leaf infection assay, 20 µL of zoospore suspension was placed on the abaxial surface of a detached leaf. Infection

was left to progress in a box lined with damp tissue paper to maintain a humid environment, at 22°C with a 16-hours photoperiod. For RNA extraction, 8 mm infected leaf discs centred on the site on inoculation were collected using a core borer. Leaf discs were flash-frozen in liquid nitrogen and kept at -80°C until RNA extraction.

For the purposes of sandwich infection assays, *P. capsici* LT1534 was grown on V8 agar. 5 mm cubes of infected V8 agar were used to inoculate 20 mL of pea broth poured into a 90 mm petri dish. The plate is sealed with parafilm was allowed to grow in the dark at 25°C for 7 days. This produced a uniform mycelial “carpet”, from which a 20 mm by 20 mm square of mycelia was cut using a scalpel and placed on the abaxial surface of a detached leaf. A second detached leaf was then placed on top, abaxial side facing the mycelia. The resultant “sandwich” was then pressed together to ensure good contact of leaves and mycelia. Infection was left to progress in a box lined with damp tissue paper to maintain a humid environment, at 22°C with a 16-hours photoperiod. For RNA extraction, the leaves were removed and discarded while the mycelia square was flash-frozen in liquid nitrogen and kept at -80°C until RNA extraction.

For the purposes of Extract and Broth inoculation experiments, plant leaf and stem tissues were collected as described. To produce Broth, 20 g of plant material was placed in 70 mL of sterile distilled water. This was autoclaved and then filtered through one layer of Miracloth. The resulting broth was then topped up to 100 mL with sterile H_2O. To produce extracts, 20 g of plant material was flash-frozen in liquid nitrogen immediately after collection. This was then ground up using a pestle and mortar and suspended in 100 mL of sterile distilled water. This was then filtered through one layer of Miracloth and topped up to 100 mL with sterile H_2O. 5 mL of 1 x 10^6 spores mL^{-1} zoospore suspension and 5 mL of either broth, extract or distilled water were mixed in

a 15 mL falcon tube. This inoculated broth, extract or water was then shaken vigorously to induce zoospore germination. The 15 mL falcon tube were stored at 22°C in the light, and placed at an angle (as close to horizontal as you dare) on its side with no lid. For RNA extraction, the liquid was discarded carefully, leaving germinated spores adhering to the side of the tube. It was flash-frozen in liquid nitrogen and kept at -80°C until RNA extraction.

3.3.2 RNA extraction

Samples were removed from the -80°C freezer and placed on ice. For broth and extract inoculation experiments 1 mL of TRIZOL reagent was added to the sample, directly in the 15 mL falcon tube. The sample was allowed to defrost on ice. During defrosting time, the sample was vortexed until it became homogenous. For leaf infection and sandwich infection assays, tissue was placed in a 2 mL Eppendorf containing a steel ball. 1 mL of TRIZOL reagent was added and this was placed in a TissueLyser II with chilled gaskets and disrupted for 1 minute at 30 Hz. The resulting lysate was centrifuged at 12,000xg for 10 minutes at 4°C to remove debris. The supernatant was removed and retained.

Following lysis in TRIZOL, 1 mL of tissue lysate from inoculation or infection experiments was then transferred to an RNase free 2 mL Eppendorf and 200 µL chloroform was added. This was then shaken well by hand until uniform and cloudy, and centrifuged from 15 minutes at 12,000 xg at 4°C. This step allows the separation of distinct phases, a bottom organic phase containing proteins, lipids and other cell debris, a middle interphase consisting of a cloudy film of DNA, and the clear upper phase which contains the RNA. The upper phase was transferred into a fresh RNase free 2 mL Eppendorf and the same volume of 70% ethanol added. The mixture was

then vortexed well, allowing the ethanol to precipitate the RNA and bind it to the column. The sample was then added to a spin column from the PureLink RNA Mini Kit (Ambion, Life technologies). The column and collection tube was then centrifuged at 12,000 xg for 30 seconds at room temperature and the flow-through discarded. The rest of the extraction was carried out as per the instructions from the PureLink RNA Mini Kit. Following elution in 40-60 µL RNase free water the Purity and quantity of RNA were assessed with nanodrop analysis. Gel electrophoresis on 1% agarose gel dissolved in Tris-Borate-EDTA (TBE) buffer was also conducted to estimate RNA integrity. TBE buffer contained 45 mM Tris-borate, 1 mM EDTA. The gel was run for 45 minutes at 75 volts RNA was then used for qRT-PCR analysis. RNA intended for sequencing underwent Qubit analysis which was conducted to better estimate RNA quantity, and RIN was found as assessed by the Agilent 2200 TapeStation systems.

3.3.3 qRT-PCR

Quantitative 2 step reverse transcriptase PCR (qRT-PCR) was conducted to quantify gene expression. Using RNA from infection and inoculation assays, RQ1 RNase-Free DNase (Promega) treatment was used (as per manufacturer protocol) to eliminate background amplification from contaminant DNA and avoid skewing the results. Following DNase, denaturation cDNA was generated using reverse transcription with SuperScript III reverse transcriptase (Invitrogen) using Oligo (DT) 20 primers as per the manufactures instructions. qRT-PCR was conducted using the real-time PCR master mix FastStart Universal Probe Master (Rox). Primer and corresponding universal probe library (UPL) probes were designed and chosen using the Universal Probe Library Assay Design Center (https://lifescience.roche.com/en_gb/brands/universal-probe-library.html). The use of the mastermix, the primers and probes, and cDNA, allowed quantification of chosen

transcripts. qRT-PCR was thus performed in StepOne™ Real-Time PCR System and data was collected and analysed using the StepOne™ software. Thereby, gene expression levels were calculated relative to expression in a germinating cyst at the initial time point (control), using the cycle threshold (ΔΔCT) methodology. Expression levels of genes were normalized to *P. capsici* ubiquitin gene (Phyca11_510705) or RPL22 gene (Phyca11_19566). Following RNA-sequencing results, minor variation in ubiquitin expression patterns meant that future gene expression quantification was normalized to RPL22.

3.3.4 RNA Sequencing (Novogene)

RNA library preparation and sequencing was conducted by Novogene. Sequencing libraries were generated using NEBNext® UltraTM RNALibrary Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations. Briefly, mRNA was enriched using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations and high temperature in NEBNext First StrandSynthesis Reaction Buffer (5X). cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-) and Polymerase I. The 3' ends of DNA fragments were adenylated, and NEBNext Adaptors with hairpin loop structures were ligated for hybridization. Fragment selection cDNA of preferentially 150~200 bp in length were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 µl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 minutes followed by 5 minutes at 95 °C before PCR. PCR amplification was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. Illumina sequencing was then conducted.

3.3.5 RNA-SEQ pre-analysis

The raw sequence reads for each sample were filtered to remove low-quality reads and sequencing of the adapters necessary for Illumina sequencing. Reads containing adaptors, reads with more than 10% undetermined bases and reads with low-quality base making up over 50% of the read, were all removed. For all samples, an average of 97.5% of total reads, deemed clean, were retained. This filtering was conducted by Novogene. Raw reads were recorded as FASTQ files, as sequencing conducted used paired ends, every sample was composed of two paired FASTQ files. Samples were uploaded to the bioinformatics platform Galaxy (https://usegalaxy.eu) and paired.

The paired reads were annotated and quantified. Annotation used filtered transcripts lists based on the *P. capsici* reference genome (LT1534 v11.0) (https://genome.jgi.doe.gov/portal/Phyca11). To conduct annotation and quantification, the Salmon quant tool (Galaxy Version 0.14.1.2) was used, to annotate and quantify transcript abundance (Patro et al., 2017).

Raw read count data pre-processing, transcript-gene mapping, and differential gene expression analysis were conducted using the 3D RNA-SEQ interface (Guo et al., 2020). In order to compare the gene expression levels between samples, the abundance of reads per transcript was normalized for transcript length and depth of sequencing using transcripts per kilobase million (TPM). TPM uses the number of reads per transcript length in kilobases and normalises it per sequence depth i.e. per million reads. Low expression transcripts were removed. As read count and the mean-variance thereof are assumed to follow a negative binomial distribution, the log2 of the read count plotted against the mean-variance allowed visual exploration of low expression transcripts (Figure 3.4). As such, counts per million reads (CPM) is set to

1 and the sample number cut offset to 12, filtering very low expressed transcripts. Thus, an expressed transcript must have ≥ 12 samples with ≥ 1 CPM expression else the transcript was filtered out. Exploratory principal component analysis (PCA) was conducted to assess the batch effect. This analysis was used to determine the variance between biological replicates, where differences could be caused by samples being prepared under differing conditions. Although the PCA showed a distinct batch effect variance in biological replicate 3, the data was not modified to remove this effect. Finally, so library composition did not sway the differential expression analysis, the data was normalised using TMM (weighted trimmed mean of M-values). All data pre-processing was carried out using the 3D RNA-seq App.

3.3.6 RNA-SEQ differential gene analysis

Differential expression analysis was conducted using the Limma-Voom pipeline/ R package via the 3D RNA-SEQ App (Guo et al., 2020). The P-value for the significance of differential expression of genes was calculated using F-test. The adjusted P-value for significance was 0.01, adjusted using the BH method.

3.4 Discussion

It is crucial for the understanding of host specificity and dynamic host adaption that organisms with broad host range such as *P. capsici* are studied whilst causing disease on different host plants. To be able to combat broad host pathogens in the field and in the multiple crop plants in which they are able to cause disease, there is a need to understand the determinants and mechanisms that are utilised to dynamically adapt to an individual host or to multiple hosts. Broad transcriptomic analysis and comparison of multiple host infections is key to increasing our understanding in this area. However, RNA sequencing of *P. capsici* infection *in vivo* has many limitations,

especially at important early time points of infection, where pathogen biomass is very low. This study demonstrates the utility of an *in vitro* methodology as a model system for transcriptomic analysis with the potential for experimental design adaptation to other oomycete pathogens and other host systems. This methodology has been utilized to examine the difference in the gene expression of *P. capsici* in two different host plants.

3.4.1 *In vitro* methodology validation

In this study, three *in vitro* methodologies were trialled for their ability to produce host dependent differential expression of the gene in a manner that had previously been observed (Unpublished Huitema Lab). Initial experiments indicated that the use of plant extracts was able to induce differential gene expression, similar to that observed in microarray data described in Chapter 1, in *P. capsici*, in all the genes tested. Of the four genes tested all showed up-regulation in CE and down-regulation in TE as predicted by previous studies (Unpublished Huitema Lab). When compared over time, culturing with plant extract gave an acceptable resemblance to plant leaf infection gene expression. However, for one of the marker genes tested, PcNPP1, the extract did not induce the expression, this gene is a marker of the late stages of infection and the shift into a necrotrophic lifestyle. It is therefore tempting to assume that extract is unable to reproduce the environmental condition that induces the switch to necrotrophy. It was concluded that this methodology has more utility in modelling the early stages of infection, especially the moment and mechanisms of host perception. Indeed, the early stages of infection are those that have been hypothesised to be key in the regulation of dynamic host adaption. The extract inoculation assay was therefore used to carry out transcriptomic analysis of the response of *P. capsici* to two host extracts (cucumber and tomato) over an 8-hour time course, in order to elucidate the

molecular determinants of the dynamic adaption of *P. capsici* to these two host plants. This methodology represents new way in which omics style studies of infectious lifecycle stages, that would otherwise be difficult or impossible to conduct, can be carried out simpley.

3.4.2 Differential gene expression in response to host may reveal host-specific dynamic adaptive mechanisms.

Previous studies have demonstrated that the molecular determinants of host specificity and range can be largely divided into two main categories; genetic plasticity and dynamic adaption. Whilst the genetic plasticity of *P. capsici* has been fairly well described (Hu et al., 2020, Lamour et al., 2012a), the ability of *P. capsici* to dynamically adapt to different host plants has to date not been analysed. In this study, it was hypothesised that dynamic adaption is largely driven by differential gene expression in response to host perception, as it has been demonstrated that many pathogens show an ability to differentially express genes in response to distinct host environments that presumably enable pathogenicity on that specific host (Harris et al., 2016, Yang et al., 2018, Kellner et al., 2014). These studies also give insight into what broad families of genes determine host pathogenicity and characterise dynamic adaptions in plant pathogens. One broad family of genes that has often been observed as being a factor in host determinism and therefore likely has a role in dynamic adaption are detoxification genes. These studies highlight genes involved in peroxidase activity, oxidoreductase activity, and antioxidant activity as important in dynamic adaption (Bowyer et al., 1995, Coleman et al., 2009, Srivastava et al., 2013, Kellner et al., 2014). More specifically the oxidoreductase (ORX1), an important antioxidant protein, has been shown to be a key gene in determining the host range

of members of the filamentous fungi genus *Fusarium* (Ma et al., 2010). GO term annotation of differentially expressed genes in this study has shown that genes involved in oxidation and reduction processes also have a role in the adaption of *P. capsici* to specific hosts, and they may have a role in the ability of *P. capsici* to infect cucumber. Additionally, transmembrane transporters, particularly ABC transporters, have often been shown to be key in resistance to certain toxins, including plant secondary metabolites and fungicides, (Yang et al., 2018, Judelson and Senthil, 2006, Pang et al., 2016). Oomycetes, particularly *Phytophthora* species show expansion and diversity within the ABC transporter superfamily, showing clear evolutionary selection and therefore a likely increase in fitness with large amounts of these types of genes (Morris and Phuntumart, 2009). ABC transporters and other related efflux pumps are clearly key to the adaption of oomycete plant pathogens to toxins and therefore novel plant environments, expression of which may be key to dynamic adaption to a specific host. Here they have also been identified, alongside other efflux pumps, transmembrane proteins to having a potential role in the dynamic adaption of *P. capsici* to its host, and to host extract.

In this study, the main trend in expression shows that genes are up-regulated on contact or within the first 2 hours of contact with CE, as well as a large up-regulation event characterising GC. This induction of genes early on in infection and unique to CE may represent a suite of differentially expressed genes with a role in perception and adaption to host. GO term annotation showed that many of these genes specifically had roles in oxidation-reduction processes. It could be suggested that these genes have an important role in detoxification of the host plant environment, are cucumber specific, and allow colonisation of the cucumber plant. An example of the detoxification role of the redox processes can be found in plant pathogens. Capsidiol,

a secondary metabolite toxin found in peppers, has been found to be detoxified via oxidation to a less toxic ketone, capsenone, during *in vitro* inoculation with fungi (Ward and Stoessl, 1972, Giannakopoulou et al., 2014). However, it is known that *P. capsici* does not detoxify capsidiol in this way, although this does not mean that other toxic secondary metabolites are not detoxified in a similar manner in *P. capsici*. These oxidation-reduction processes may also have a role in defence against ROS, a common component of the plant immune system. Interestingly, the other large family of genes that were uniquely found up-regulated in CE when compared to GC or when comparing TE and CE directly, were transmembrane transporters. It has previously been demonstrated that ABC transporters and other transporters are key in protection against toxins. Moreover, with genomic expansion of ABC transporters with respect to other *Phytophthora* species, *P. capsici* may rely on these sorts of efflux pumps to overcome sensitivity to capsidiol, and transporters may be key determinants of host range (Giannakopoulou et al., 2014). The results of this study also suggest that efflux pumps or transmembrane transports may be a key defence mechanism to hostile plant environments and therefore a determinant of host range. The ability of *P. capsici* to infect and cause disease on cucumber may rely on its ability to pump out innate secondary metabolites found in cucumber plants.

In contrast to the up-regulation event in CE, the response in TE is a down-regulation event. Few genes show differential up-regulation in TE and those that do are frequently shared between both host extracts and up-regulated when compared to GC and to later stages of the extract colonisation. This could be for several reasons, presumably tomato plants are not “easier” to infect and thus rely on fewer genes. One other possibility is that the lab strain of *P. capsici* used to inoculate the extract has been routinely grown on V8 media. The principal component of V8 media tomato fruit

juice, perhaps the lab strain of *P. capsici* has constitutive adaption to the tomato host environment and limited changes are necessary when adapting to the TE environment. Even perhaps demonstrating the importance of epigenetic factors to dynamic host adaption. Additionally, in previous studies (Unpublished Huitema Lab) where tomato infection of *P. capsici* was compared to three other host plants, tomato infection was also characterised by less up-regulating of genes (91 of 2452), and cucumber infection showed the largest up-regulation event of the 4 host plants tested (742 of 2452). Interestingly, the genes that this study showed to be up-regulated in TE and also up-regulated in CE in comparison to GC had a different GO term composition. The largest group of genes, in this case, being phosphorylation, translation, protein folding and transport. This suggests a change in cell signalling and a move to a longer-term change in cellular make up. It may represent a more sustained adaptive mechanism initiated by the initial perception and dynamic reactionary phase, perhaps representing a shared pathway common to adaption to both hosts.

The vast majority of differentially expressed genes in all comparison groups, but especially those that had been found uniquely up-regulated early in response to CE, are integral parts of the cell membrane. This demonstrates that the majority of differential expression is causing a large change in protein complement at the membrane. Many of these will be transporters as already mentioned, however, the membrane may have other key functions during host perception and host colonization, acting as a signalling hub for dynamic adaption. The host perception signalling mechanisms of plant pathogens are not well studied, but reports have suggested that the surface of the pathogens has a role. It has been demonstrated that both physical and chemical signals can induce a response in plant pathogens (Kou and Naqvi, 2016). As in most species, phosphorylation is a key part of this signalling, and MAP

kinase pathways and other phosphorylation based signalling pathways have been described in plant pathogens used to respond to host signals (Kou and Naqvi, 2016, Liu et al., 2011). Proteomic studies of the pathogen membrane reveal it as a key area for many functions related to pathogenicity and adaption to host, including metabolism, oxidative stress response, autophagy and cell death. Many of which are controlled by changes in phosphorylation that occur at the membrane (Escobar-Niño et al., 2019). This study and others reveal the importance of the membrane as a site of host perception and dynamic adaption.

3.4.3 Marker gene expression and house-keeping gene expression reveals aspects of in vitro phenotype.

Expression profiles of marker genes gave insight into the similarity of the *in vitro* methodology to leaf infection, and the lifecycle stages perhaps replicated in extract inoculation. Neither PcHMP1 nor PcNMRA1 were induced in either extract, however, PcNMRA1 was induced in GC. PcNMRA1 controls nitrogen metabolism and may be induced to increase nitrogen recycling during metabolite shortages. This may demonstrate that whilst GC may have some starvation phenotypes, metabolite sensing and nutrient extraction is occurring in both extracts, suppressing PcNMRA1 expression (Pham et al., 2018). PcCDC14 is a known marker of sporulation and shows a marked decrease in expression as the inoculation progresses as *P. capsici* moves out of the spore stages into a more mature phase. This change in PcCDC14 would also be observed in leaf infection.

Genes that were known to be differentially expressed, specifically genes that were known to be up-regulated in cucumber at 16 hpi in infected leaf tissue (Unpublished Huitema Lab) were also analysed. Although the same expression profile was observed, in that they were up-regulated in cucumber in comparison to TE, they were

not significantly DE. However different stages of infection were being compared, and it is likely that the later stages of infection are not replicated during *in vitro* culture in extract. There is no information or studies on what would be expected to be differentially expressed during the early stages of infection in *P. capsici* for comparison.

The final set of genes that the expression profile was analysed are the housekeeping genes. These genes, in theory, should be uniformly expressed across all time points. However, this analysis shows a large variation in expression across the different condition. This is particularly worrisome as these genes are frequently used as a baseline level of expression for relative expression calculation such as the ΔΔCT method used for qRT-PCR analysis. In light of this, other genes that have been identified in other studies in the Huitema lab were tested for expression, genes that are frequently used as a positive control during routine PCR experiments. The large ribosomal subunits (RPLs) were examined for expression levels with the expectation of high and stable expression in all conditions. They showed as expected, high levels of expression and relative to the other housekeeping genes tested, were very similar in expression levels. Going forward these RPLs and in particular, RPL22 which showed the most stable expression of the four genes tested, will be used as a routine housekeeping gene for relative expression analysis as in qRT-PCR. It should also be noted that caution should be practised in assuming a stable expression of housekeeping genes even if routinely used in the lab and literature.

3.4.4 RNA-SEQ validation.

In order to validate the RNA-SEQ experiment, several genes were chosen for evaluation of their expression profile via qRT-PCR. Due to time constraints, only six

genes were tested. Of the six genes tested, none showed exactly the same expression profile as the RNA-SEQ predicted. However, in two of the genes, the host differences in expression were maintained. Expression in the leaf infection assay was analysed at the same time points and two of the genes displayed the same host specific differential expression as observed in the RNA-SEQ. A caveat lies over these results due to the lack of experimental replicates. The results do suggest that the methodology is highly variable and it may be that more replicates are necessary to effectively predict DE genes.

3.4.5 Conclusion and further work

This study has developed a novel methodology for *in vitro* infection assessment of *P. capsici* that has the potential to be utilised with other oomycete plant pathogens and with multiple host plants enabling broad transcriptomic analysis. This novel methodology has been used for a two-host comparison and has shown that extract is able to induce host-specific differential expression. In general, the response to cucumber extract is characterised by an early up-regulating event of genes with putative roles in detoxification and the efflux of toxins. The response to tomato extract is characterised by down-regulation, however, those genes that are up-regulated when compared to germinating cysts (control) are also up-regulated in cucumber extract and are largely characterised by genes that have putative roles in cell signalling and translation, pointing to a long-term dynamic adaptive response found in both host extracts. Although a question still exists over the reproducibility of this methodology and the predictive ability to generate genes with key roles in host specificity, potential key insights into the mechanisms that define dynamic adaption have been gained. Furthermore, the role of many genes identified should be examined and the

mechanisms identified considered in further studies. Characterisation of the genes identified here during leaf infection and their effect on the pathogenicity of *P. capsici* in different hosts is a key step towards understating how dynamic adaption drives host specificity and host range.

Chapter 4: Proteomic analysis of *P. capsici* during an *in vivo* model of infection of two distict hosts

4.1 Introduction

Few studies have taken a broad 'omics' style analysis of the effect that a host plant has on a plant pathogen during infection. Especially concerning the analysis of the mechanisms and molecular basis of pathogen host range. The broad host ranged pathogen *Phytophthora capsici,* can infect multiple host plants, but how it maintains this broad host range is still vague. However, as described in previous chapters, there is evidence of differential gene expression as a driver of dynamic adaption to the host environment (Harris et al., 2016, Allan et al., 2019, Yang et al., 2018, Ma et al., 2010). As such, joint transcriptomic and proteomic analysis of the effect of multiple hosts on the pathogen may shed some light on the deeper cellular mechanisms, and molecules with roles in broad host range pathogenicity in oomycetes and other broad host ranged plant pathogens. 'Omic' analysis of various lifecycle stages in *Phytophthora* can also be useful in highlighting key proteins and their potential life stage-specific functionality (Resjö et al., 2017)

Proteomics is an indispensable tool when characterising the protein content of a cell. As such it has long been used as an unbiased approach to understand the factors that contribute to plant immunity and pathogen virulence and their interplay. Tracking changes to a pathogen or plant cells through the infection lifecycle, from PAMP mediated recognition and immunity (PTI) through to effector release, and the susceptibility or immunity that is triggered from those effectors (ETI/ ETS), allows identification of the key molecular players in those battlegrounds. As such, many

studies have been conducted *in planta* investigating the plant's response to assault with a pathogen. (Xiao et al., 2019, Larsen et al., 2016). Proteomics can also find utility in probing other areas where pathogen plant interactions are taking place. For example, the apoplastic space is a key compartment in which plant-microbe conflicts play out and proteomics can easily detect associated signatures (Delaunois et al., 2014). Additional studies have look at a particular pathogens response to chemicals. For example, the response of the oomycete pathogen *P. capsici* to a fungicide flumorph. Identifying 181 proteins considered to contribute to the adaptive response of *P. capsici* to such a fungicide. (Pang et al., 2016). Studies have also examined changes in *Phytophthora* proteome during its life cycle, with an emphasis on pre-infectious lifecycle stages (Resjö et al., 2017, Pang et al., 2017).

However, as these studies increase our picture of how the plant responds to a pathogen and how a pathogen can respond to its environment, there is still a gap in our understanding of plant-pathogen interactions. It is more challenging to investigate the response of the pathogen to the perception of a potential host *in planta*, than that of studying the host's response. Proteomic *in planta* studies of a pathogens response to host would be near impossible to conduct, as stated in previous chapters, the perineal problem is a great excess of plant material in any potential sample blocking any possible signal from pathogen molecules. Especially when considering desired breadth and depth in signal strength to make any differential expression analysis possible. Therefore, other non-*in planta* methods of deriving sample for proteomic analysis are necessary. The work conducted in Chapter 3 has validated and conducted similar ‘omics’ based experiments. In order to investigate the changes that a pathogen undergo during infection, it was necessary to develop a methodology to isolate the pathogen material from the plant. Chapter 3 details the development and

validation of an extract inoculation assay, a system in which the environment to which *P. capsici* is exposed during the early stages of infection is mimicked. This methodology was used to conduct a transcriptomic investigation of the effects of different host extracts on the early stages of *P. capsici* adaptation to its host. This approach enables us to attempt to patch the gap left by previous studies; *Phytophthora* life cycle analysis, which has invariably focused on the non-infectious life cycle stages, and in studies of plant-pathogen interactions which often focus on the changes in the plant's system and not the pathogen response.

RNA sequencing of the extract inoculation assay demonstrated that challenging *P. capsici* with different host plants affects the genes being expressed, and is therefore influenced by its host environment and its resident signals.

For the experiment that was carried out in Chapter 3 the effect and difference between only two host extracts was investigated (Tomato and Cucumber). RNA sequencing of this extract inoculation assay conducted with two host plants, revealed initially that incubation of the *P. capsici* zoospores in host extracts induced a response, more specifically it was found that there is a large quantity of differential expressed genes when comparing between the hosts at 2 hours post inoculation (hpi). It highlighted genes involved in oxidation and reduction processes, membrane proteins, especially transmembrane transporters, particularly ABC transporters. It also revealed a large number of proteins localised to the membrane, and proteins that have phosphorylation process to be important in host perception or dynamic adaption. The results of the RNA sequencing demonstrated the potential utility and simplicity of the extract inoculation methodology in 'omics' analysis of a pathogen in response to its host. To both further validate this methodology and the data generated by the RNA sequencing, a dual 'omics' approach was deemed to be valuable as transcript levels may not

always represent protein levels and their predicted impact on *P. capsici* phenotype. Therefore a separate proteomic analysis of the two host extract inoculation experiment was also conducted.

4.2 Results

4.2.1 Protein extraction and quality control

Proteins were extracted using multiple methodologies to see which would perform the best in term of total concentration and integrity as viewed on Coomassie stained gels. Eight methodologies were trialled, those were a combination of two different extraction buffers GTEN-T and GTEN-DM, two different lysis aids: vortexing or using the TIssueLyses II machine, and then whether the protein was purified or not (Figure 4.1). By simply examining visualised total protein extraction on the SDS polyacrylamide gel, it was determined that GTEN-DM gave a better concentration and that vortexing or using the TissueLyser II had little to no effect on the integrity of total amount of proteins isolated, nor did whether the proteins were methanol precipitated to purify or left in extraction buffer (Figure 4.1). Therefore GTEN-DM buffer was used for total protein extraction, the tissue was then lysed only through the aid of vortexing to mix thoroughly. Due to the contaminating nature of the high concentration and mix of detergents present in GTEN-DM, and the fact that it appeared to have little effect on the amount of total protein recovered, the protein was purified using methanol precipitation, prior to suspension in a Tris-SDS buffer for proteomic analysis.

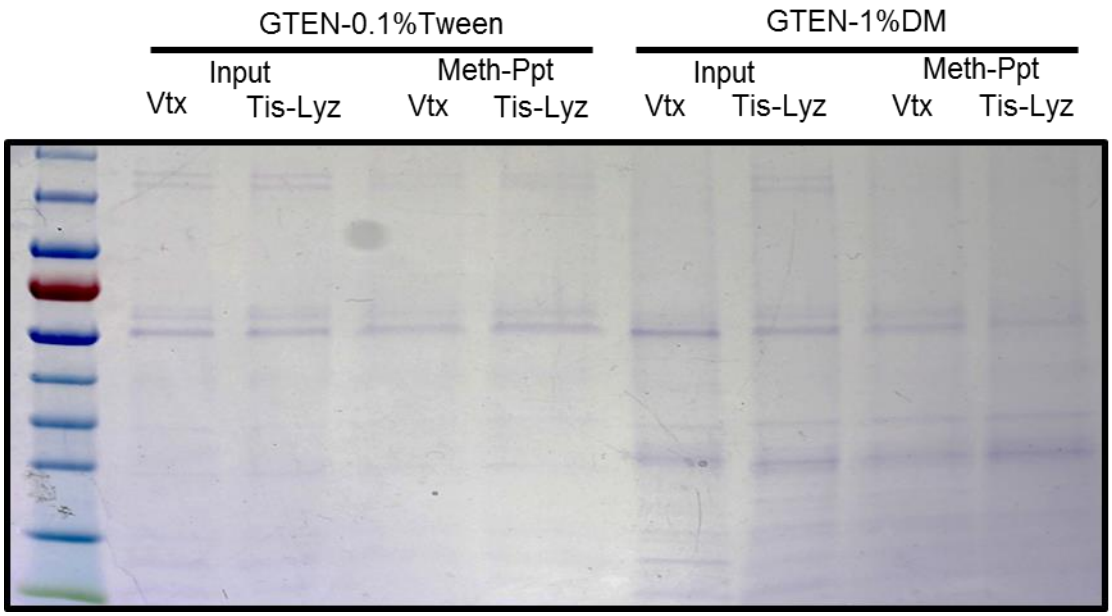


Figure 4.1: coomassie stained SDS-page gel or protein extraction methodologies.

This shows the 8 proteins extraction methodologies trailed, with the two different buffers (GTEN-0.1%Tween and GTEN-1%DM), either methanol purified (Meth-ppt) or non purified (input), and using the an vortex or a TissueLyser II to aid lysis. Sample were then run on an SDS-page gel and stained with coomassie to visualise. The protein ladder is PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa (thermofisher).

4.2.2 Extract inoculation and differentially abundant protein identification

Following *P. capsici* zoospore isolation the zoospores were subjected to seven different conditions. They were either challenged with cucumber extract (CE) or tomato extract (TE), samples were then collected at 2 hours post inoculation (hpi), 4 hpi, or 8 hpi, giving six possible conditions. The final seventh conditions was zoospores left in water to germinate also known as germinating cysts (GC) and collected at 2 hpi. This resulted in a total of seven conditions CE 2hpi, CE 4 hpi, CE 8 hpi, TE 2 hpi, TE 4 hpi, TE 8 hpi, and GC 2 hpi. For each condition, four biological replicates were generated and total protein extracted. Label-free quantitative Mass spectrometry analysis was conducted. To identify proteins that are present in each of the samples the peptides were searched against the *P. capsici* predicted proteome using the Maxquant software

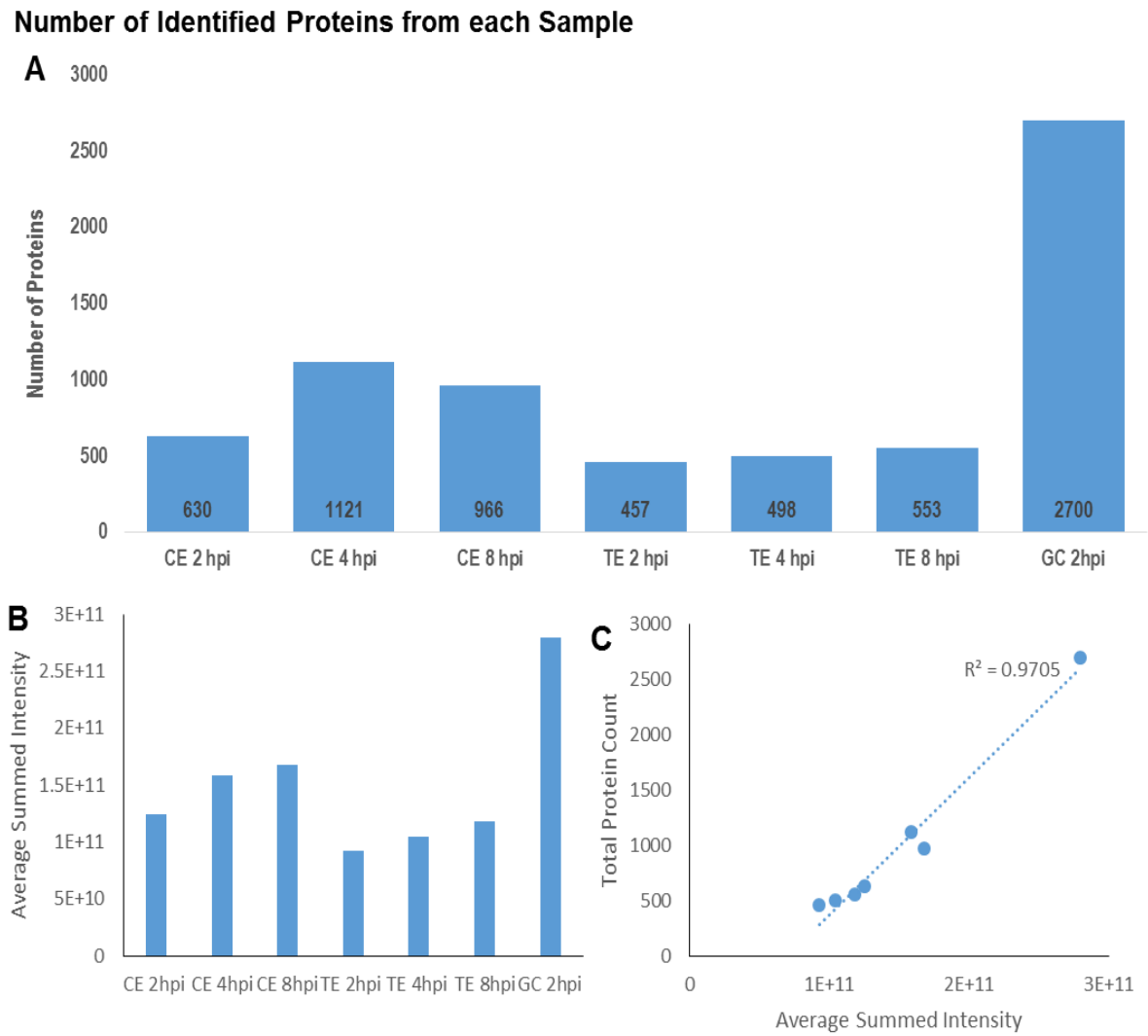


Figure 4.2: Number of Proteins Identified in each sample.

For each of the 7 condition, the number of proteins that were identified in at least 1 of the four replicates is shown (A). In addition the average intensity for each protein identified was summed between the 4 samples and the average calculated. The total intensity for each condtion averaged for each protein as described above was then summed (B). The linear relationship between average summed intensity and the total number of proteins for each condition was calculated, with a pearsons correlation (R) = 0.9705,

(Cox et al., 2009). 2914 proteins were identified of a potential 19,805 from the current best-predicted protein database for *P. capsici*. By far the most proteins identified were in the GC sample, 2700 (92.66%) proteins were found in at least one of the four GC replicates. Of the samples challenged with host extracts (TE/CE) those derived from the challenge with CE had the most identified proteins, with 630, 1121, and 966 identified respectively from CE 2 hpi, CE 4 hpi, and CE 8 hpi. Searches in samples

from the TE inoculation identified fewer proteins with 457, 498, and 553 identified respectively from TE 2 hpi, TE 4 hpi, and TE 8 hpi. (Figure 4.2 A).

For each condition, there were 4 biological replicates, resulting in 28 total samples. The total intensity of all identified proteins in each sample was summed for each condition and the averaged summed intensity across the replicates was calculated to estimate relative protein abundance and potential plant protein contamination. A pattern of average summed intensity was calculated for each condition that was similar to the number of proteins identified in each condition (Figure 4.2 B). Plotting the total protein count against the average summed intensity for each condition shows a close linear relationship (R^2=0.9705) between the two (Figure 4.2 C).

To examine the effect of host extract on *P. capsici* and the difference in protein intensity induced by the two hosts extracts, multiple comparisons of the seven samples were made. For the purpose of these results, the comparisons have been separated into two distinct groups, those between the both host extracts and GC; Extract versus GC, and those between the two host extracts; CE versus TE. For the differential protein intensity analysis that is shown here, differential analysis was conducted for every protein that was identified in any sample. However statistical significance test could only be conducted if a minimum of two of the replicate samples from both the conditions compared had intensity data. As such the differential protein set that has been analysed here are for those proteins that had intensity data from as a minimum 2 of the 4 replicates. For completion and validation the number of differentially abundant proteins has been calculated using more stringent requirements for the data. Using as a minimum 3 of the 4 replicates having intensity data, or all 4 replicates having data, each reduced the number of statistically significant proteins with differential abundance (Figure 4.3). Showing that a lot of those proteins that were

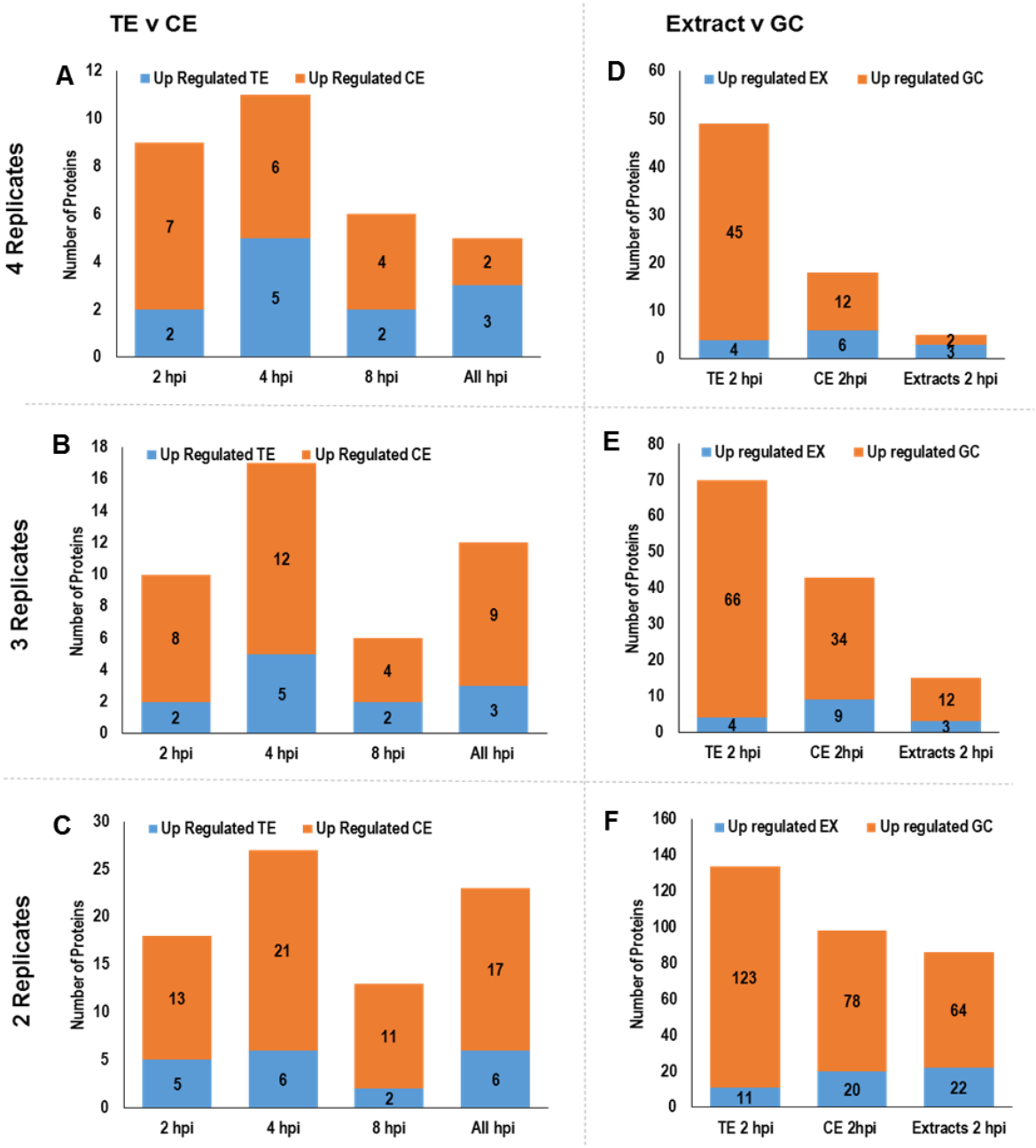


Figure 4.3: Comparison of statistical analysis of differential protein intensity.

The number of proteins with differential intensity from each 7 comparisons. These seven comparisons are separated into four Tomato Extract (TE) and Cucumber Extract (CE) comparisons at 2 hours post inoculation (hpi), at 4 hpi, at 8 hpi and at All hpi pooled. And three Extracts (EX) and germinating cycst (GC), TE compared with CG at 2hpi, CE compared with GC at 2hpi and the two Extracts pooled at 2 hpi compared with GC at 2 hpi. For all seven comparison, the number of proteins with differential intensity are shown if the comparisons are only carried out if there is data present for that protein for a certain number of replicates in each sample. The number of replicates with data necessary for comparison of a proteins intensity to be carried out is on the right A and D needed 4 replicates, B and E needed 3 replicates, and C and F needed 2 replicates. The number of proteins upregulated in CE (A,B,C) and GC (D, E,F) is shown in orange, the number of proteins upregulated in TE (A,B,C) and in Extract (D,E,F) is shown in blue. The number of those proteins is shown in each bar.

found to have differential abundance were missing in one or two of the replicates. However, due to the variability of the data, and the fact that many interesting proteins may have low abundance in some of the conditions or samples, or indeed may have been out competed by plant material, it is our view that having at least 2 replicates with data is sufficient for this analysis. Indeed, the pattern of the number of differentially abundant proteins between comparisons remains the same. It is important to note, however, if any of these proteins are to be explored further, the pattern of abundance should be fully validated. If this analysis was conducted again, it would be suggested to only use proteins present in at least 3 replicates. The difference in the numbers of differentially abundant proteins found is greater between analysis using 2 – 3 as a minimum number of replicates where proteins were present, than the difference between 3 – 4 (Figure 4.3).

4.2.3 Extracts compared with germinating cysts: Host extract induced differential protein abundance

Initial analysis to examine whether host extracts induce differences in the proteome was conducted.For this three comparisons were made: CE 2 hpi versus GC 2 hpi, TE 2 hpi versus GC 2hpi, and the 2 hpi time point with both CE and TE pooled versus GC 2 hpi (Extract vs GC 2 hpi). Volcano plots of these comparisons were generated

(Figure 4.4 A-C). In both the CE 2 hpi and TE 2 hpi compared with GC 2 hpi, a increase in abundance of a number proteins can be seen in the GC condition (Figure 4.4 A-B). This big shift of gene on the left can also be seen when comparing Extract at 2 hpi with GC (Figure 4.4 C). From all these proteins, those that were significantly

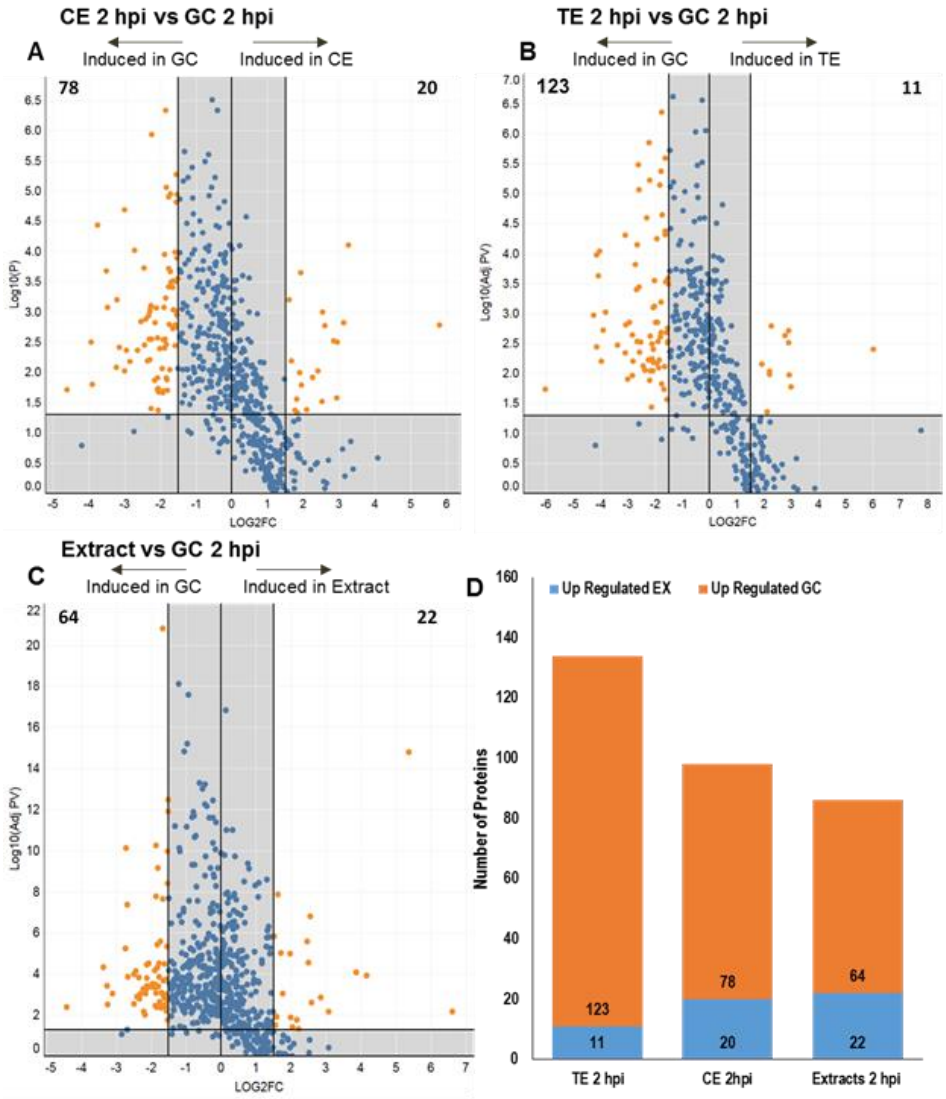


Figure 4.4: Volcano Plots showing proteins intensity for comparisons of Host extracts and GC.

For three of the comparisons (CE 2hpi vs GC 2hpi, TE 2hpi vs GC 2hpi and the pooled condition of CE and TE 2hpi here name Extract vs GC 2hpi) the LOG2 fold change (FC) and the adjusted P value (from a two tail t-test with Bonferroni correction) for each protein found in at least 2 replicate form the two compared conditions was calculated. The LOG10 p value and the LOG2FC for each protein for each comparison have been plotted. Those proteins that have been shifted to the right are comparably induced in extract (CE, TE, or pooled) and those proteins that have shifted to the left are induced in GC. Those significantly differentially abundant proteins i.e. with a Log2FC ≤-1.5 or ≥1.5, and a -log10 P value ≥ 1.3 have been shown in orange, and the point where a protein is significant has been indicated with a black line. Protein in the greyed area represented by a blue dot are not significantly differentially abundant (A-C) The number of significantly induced proteins in each subset are shown at the top of each plot.. The number of significantly differentially abundant proteins in each condition in comparison to GC 2hpi (TE 2hpi, CE 2hpi, and Extracts have been shown with those up-regulated in extract (CE, TE or pooled) shown in blue in the bottom portion of the bar, and those up-regulated in GC shown in orange in the top portion of the bar (D).

differentially abundance where isolated (Log2FC ≤-1.5 or ≥1.5, and -log10 P-value ≥ 1.3). For all three comparisons between extracts and GC, the volcano plots show a increase in abundance of a greater number proteins in GC than in extracts, which was borne out by the number of significantly differentially abundant proteins. 123 and 78 proteins had greater abundance in GC that TE and CE respectively. With only 11 proteins in the TE condition and 20 in the CE condition having greater abundance in those respective extracts. Also, when all extract samples were pooled, this gave 64 proteins with greater abundance in GC and 22 with greater abundance in extracts (Figure 4.4 D).

Examining the overlap between the collections of differentially abundant proteins can give clues to how the response of *P. capsici* to different extracts can differ (Figure 4.5).

There is a large overlap, 78 proteins, which can be found with greater abundance in GC when comparing to both host extracts (Figure 4.5 B). However, of those proteins found with greater abundance uniquely in GC when comparing to a single host extract, all 45 are found when comparing to TE, there are no proteins found with greater abundance in GC compared to CE that are not also found in the comparison with the other host extract. Or to put it another way, every protein that has a higher abundancein GC than in extract is also found with increased intensity in CE (Figure 4.5 B). This overarching pattern of abundance was shown more clearly by plotting the intensity of all proteins which were found with greater abundance in GC and not in *P. capsici* in host extract on a heat map. It can clearly be noted that in general those proteins that have high abundancy in GC also have higher abundance in CE, especially CE 4 hpi and 8 hpi, relative to their intensity in TE samples (Figure 4.6).

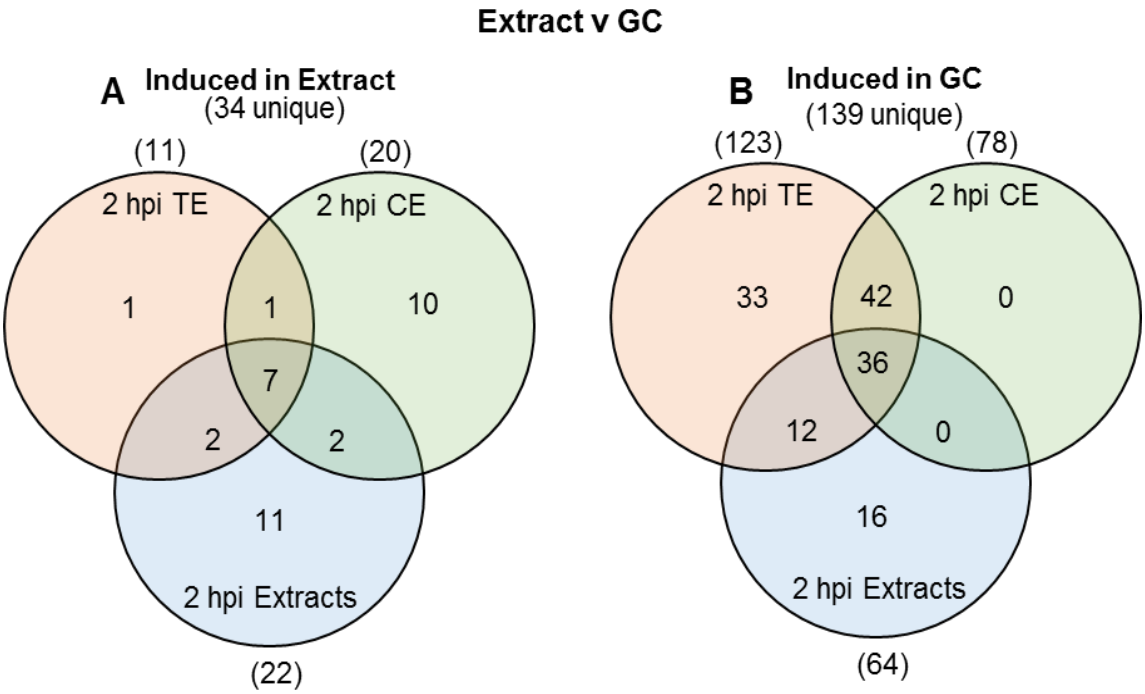


Figure 4.5: Venn diagrams showing the overlap between those proteins induced in extracts, and the overlap of those proteins induced in GC.

The overlap in the proteins that were shown to be significantly abundant in extracts (CE 2hpi , TE 2hpi and pooled (extract 2hpi) compared to GC (A). And the overlap in the proteins that were shown to be significantly abundant in GC compared to extracts (CE 2hpi , TE 2hpi and pooled (extract 2hpi) (B). The total number of proteins in each comparison is shown in brakets by the circle and the total number of unique proteins in the three comparisons together is shown in brackets under the title.

Although less than the number proteins with greater abundance in GC, some proteins were found with increased abundance in host extracts (2 hpi) compared to GC. 34 unique proteins were significantly abundant in *P. capsici* in host extracts. Of these proteins, the majority were found with increased intensity in CE 2 hpi, 20, whilst only 11 were found with increased intensity in TE 2 hpi. Pooling these two conditions and time points gave an additional 11 uniquely significantly differentially abundant proteins (Figure 4.4-4). Again, the overlapping abundance of these proteins can reveal clues to the response of *P. capsici* to extract. Few of these significantly differentially abundant proteins, only 3, are uniquely found with increased

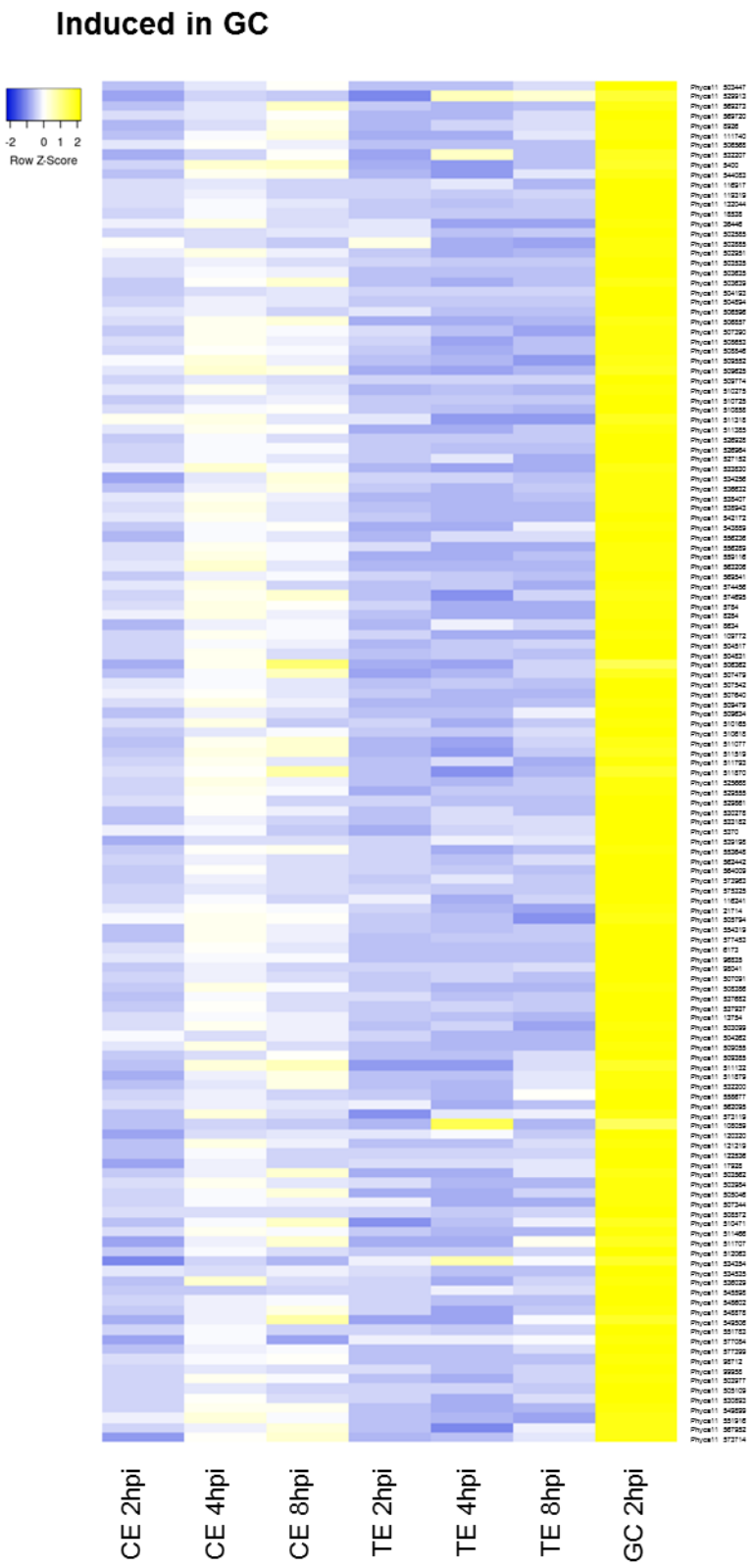


Figure 4.6: **Heat map of 139 proteins with greater abundance in GC compared to EX.**

Expression is shown by adjusted z-score. Down regulation is shown in blue and up regulation is shown in yellow.

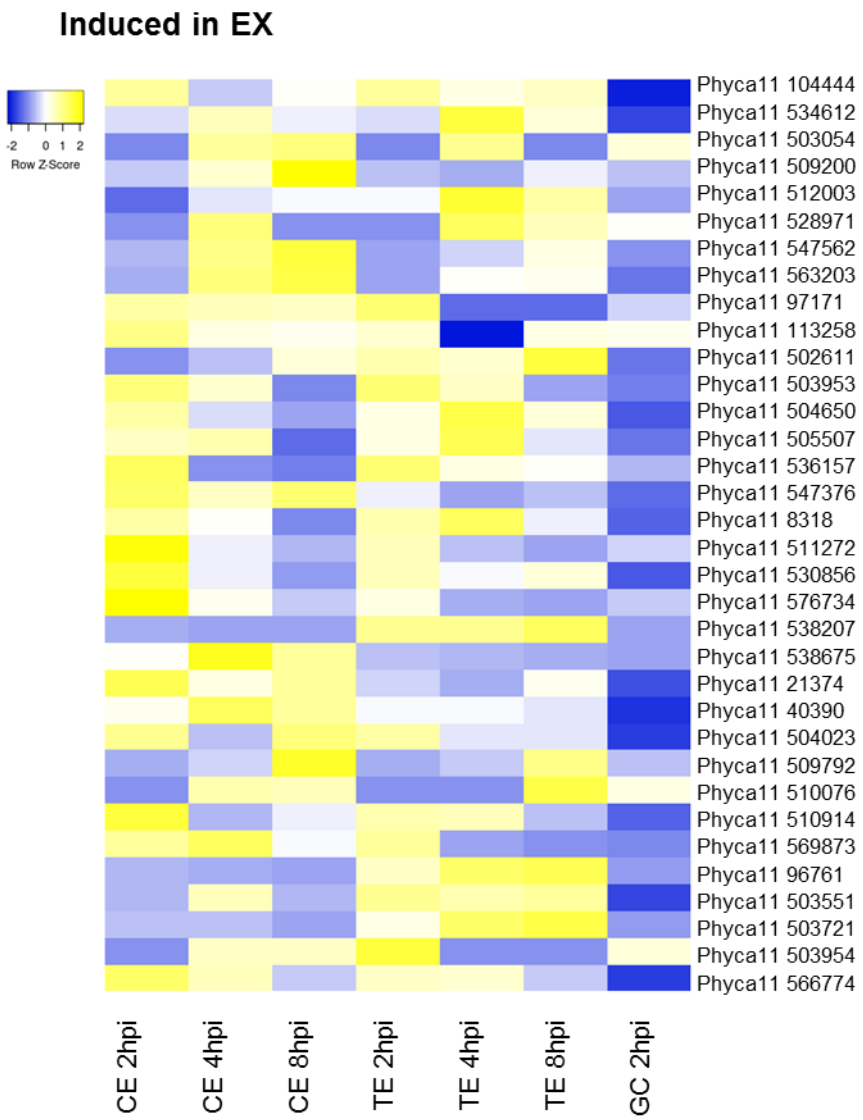


Figure 4.7: Heat map of 34 proteins with greater abundance in EX compared to GC.

Intesity is shown by adjusted z-score. Down regulation is shown in blue and up regulation is shown in yellow**.**

intensity in TE 2 hpi. Whilst, 8 proteins were found in both host extracts, and the majority (12) of these proteins with greater abundance can be found uniquely in CE 2hpi (Figure 4.5 A). This can, perhaps, be seen more clearly in the heat map, where it can be seen that abundance levels at 2 hpi can be separated up into two major categories, those proteins that have higher abundance in both extracts (CE 2hpi and

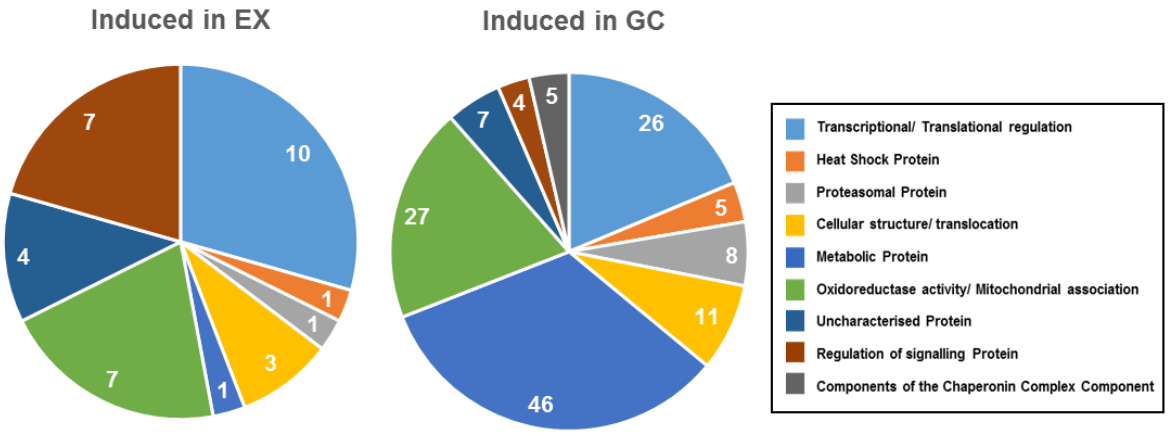


Figure 4.8: Distribution of broad functionality categories in proteins with significantly differential abundance between extracts and GC.

Using GO terms and KOG annotation each protein was manually assigned to a broad functionality group that was enriched in the differentially regulated proteins. These consist of 139 proteins up-regulated in GC and 34 proteins up-regulated in extract. From the top of both charts moving clockwise those categories are Transcriptional/ Translational regulation, Heat Shock Proteins, Proteasomal Proteins, Proteins involved in Cellular structure or translocation, Metabolic Proteins, Proteins with Oxidoreductase activity or Mitochondrial association, Uncharacterised Proteins, Regulation of signalling Proteins, and Components of the Chaperonin Complex. Note that of those proteins induced in Extract none are Components of the Chaperonin Complex.

TE 2hpi) and those that are found only higher in just CE 2hpi. Very few proteins can be found uniquely abundant at TE 2hpi (Figure 4.7).

4.2.4 Extracts compared with germinating cysts: Protein Functionality and intensity profile

To examine the potential role of these proteins, Biological process GO terms, and EuKaryotic Orthologous Groups or KOG annotations were found for each. Informed by these results the proteins were then manually subdivided into 9 different categories. The categories were, 1) transcriptional / translation regulation, 2) Heat Shock Proteins, 3) Proteasomal, 4) Cellular structure / translocation, 5) Regulation/ signalling, 6) Chaperonin complex component, 7) Metabolic, 8) oxidoreductase activity/

mitochondrial, and 9) Uncharacterised (Figure 4.8). Described below, for each these protein groups in turn, is an examination of their relative differential abundance in GC and extract, and their intensity profile across the time course of the experiment.

Transcriptional regulators and translational regulators are highly abundant in both GC and *P. capsici* in host extracts. 26 (18.7%) of these proteins are found with increased intensity in GC and 10 (29.4%) in extract samples (Figure 4.8). The GO terms and KOG annotation for these proteins vary. However, for those with higher intensity in GC, there is a high abundance of KOG annotated tRNA synthases, in fact, 13 of the 26 proteins in this category are various tRNA synthases (Supplementary Table 3). Whereas those proteins found with increased intensity in response to extracts, 5 of the 10, are various ribosomal proteins (Supplementary Table 3). The intensity profile of those transcriptional, translational regulators with higher intensity in GC, frequently increased in intensity as the time course progresses in the TE and CE samples (Figure 4.9 A). The same pattern cannot be seen in those proteins that have increased intensity in extracts, most seem to have a stable abundance over the time course of the experiment or have a small decrease (Figure 4.9 B). Notable of these proteins is Phyca11_538207, with a KOG annotation of “DEAH-box RNA helicase” with a particularly high intensity in TE at all hpi (Figure 4.9 B).

Amongst the differentially abundant proteins Heats shock proteins (HSP) or Molecular chaperones HSPs, were frequently identified by KOG annotation. They also have relatively similar abundance, 5 (3.6%) with greater abundance in GC and 1 (2.9%) with greater abundance in extract samples (Figure 4.8). Similar to the proteins identified as transcriptional/translational regulator, those proteins with increased intensity in GC frequently a show progressive increase in intensity in extracts as the time course of the experiment progresses (Figure 4.9 C). In addition, the single Molecular

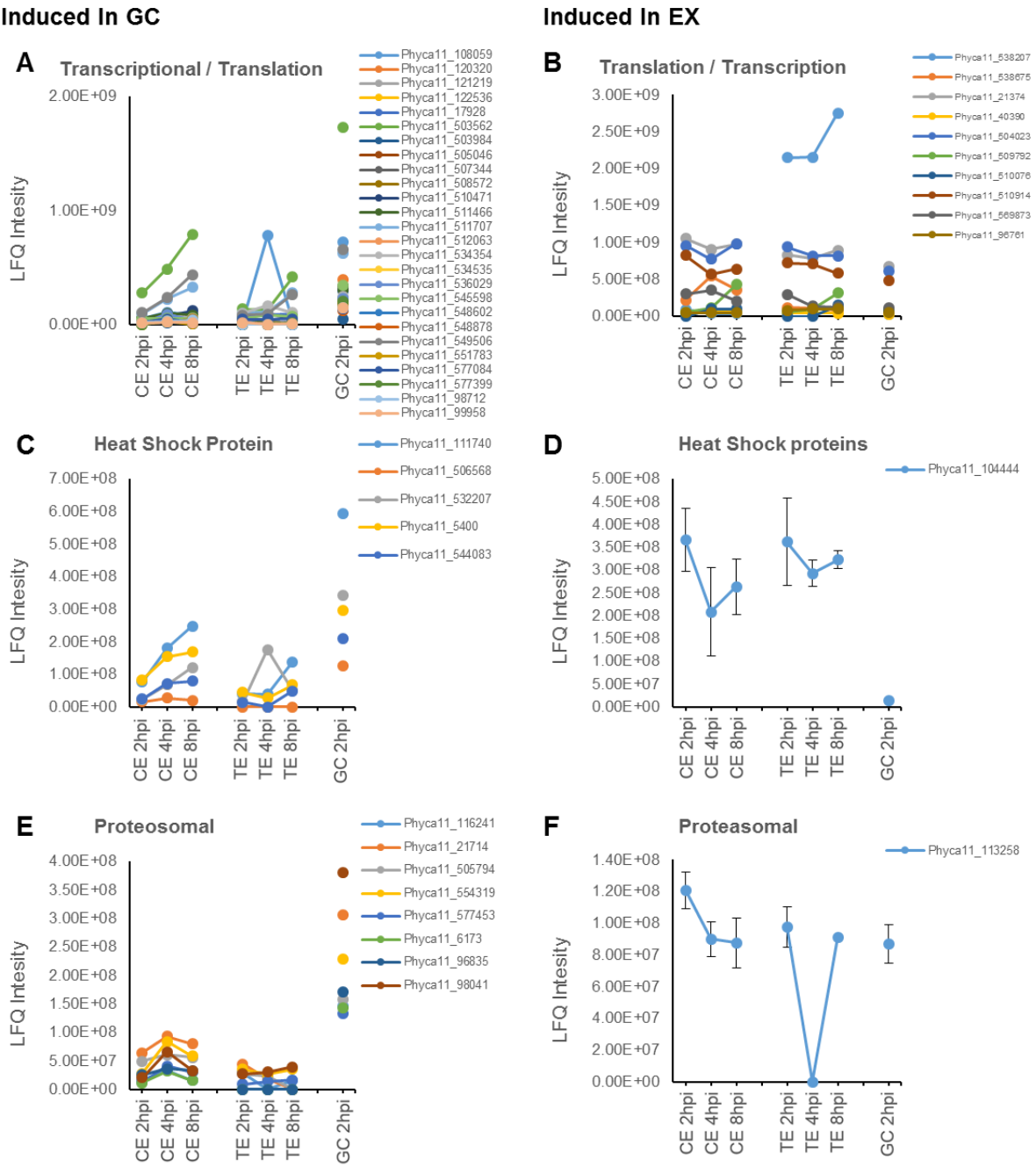


Figure 4.9: Intensity profile of Proteins.

For each protein that was manually annotated with the broad functionality category Transcriptional/ Translational regulation (A,B), Heat Shock Proteins (C,D), Proteasomal Proteins (E,F), the average intensity from all replicates for each condition (CE 2, 4, 8 hpi, TE 2, 4, 8 hpi, and GC 2 hpi) has been plotted. The protein from these functionality categories that were induced in GC (A,C,E) and induced in Extract (B,D,F) are shown here. If there are fewer than 4 proteins plotted on a single graph the standard deviation has been shown with error bars.

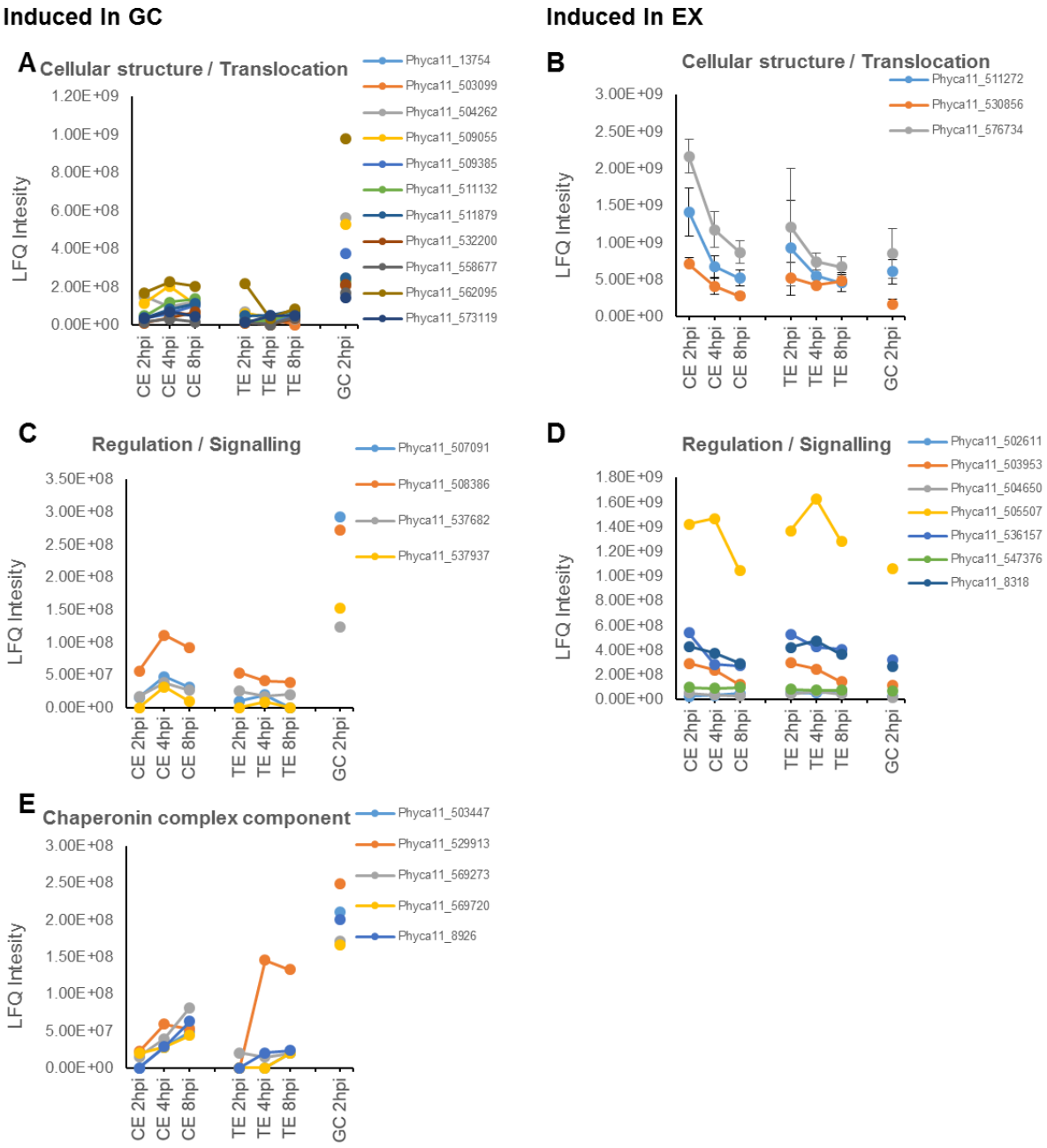


Figure 4.10: Intensity profile of Proteins.

For each protein that was manually annotated with the broad functionality category Cellular structure or translocation (A,B), Regulation of signalling Proteins (C,D), and Components of the Chaperonin Complex (E), the average intensity from all replicates for each condition (CE 2, 4, 8 hpi, TE 2, 4, 8 hpi, and GC 2 hpi) has been plotted. The protein from these functionality categories that were induced in GC (A,C,E) and induced in Extract (B,D) are shown here. If there are fewer than 4 proteins plotted on a single graph the standard deviation has been shown with error bars.

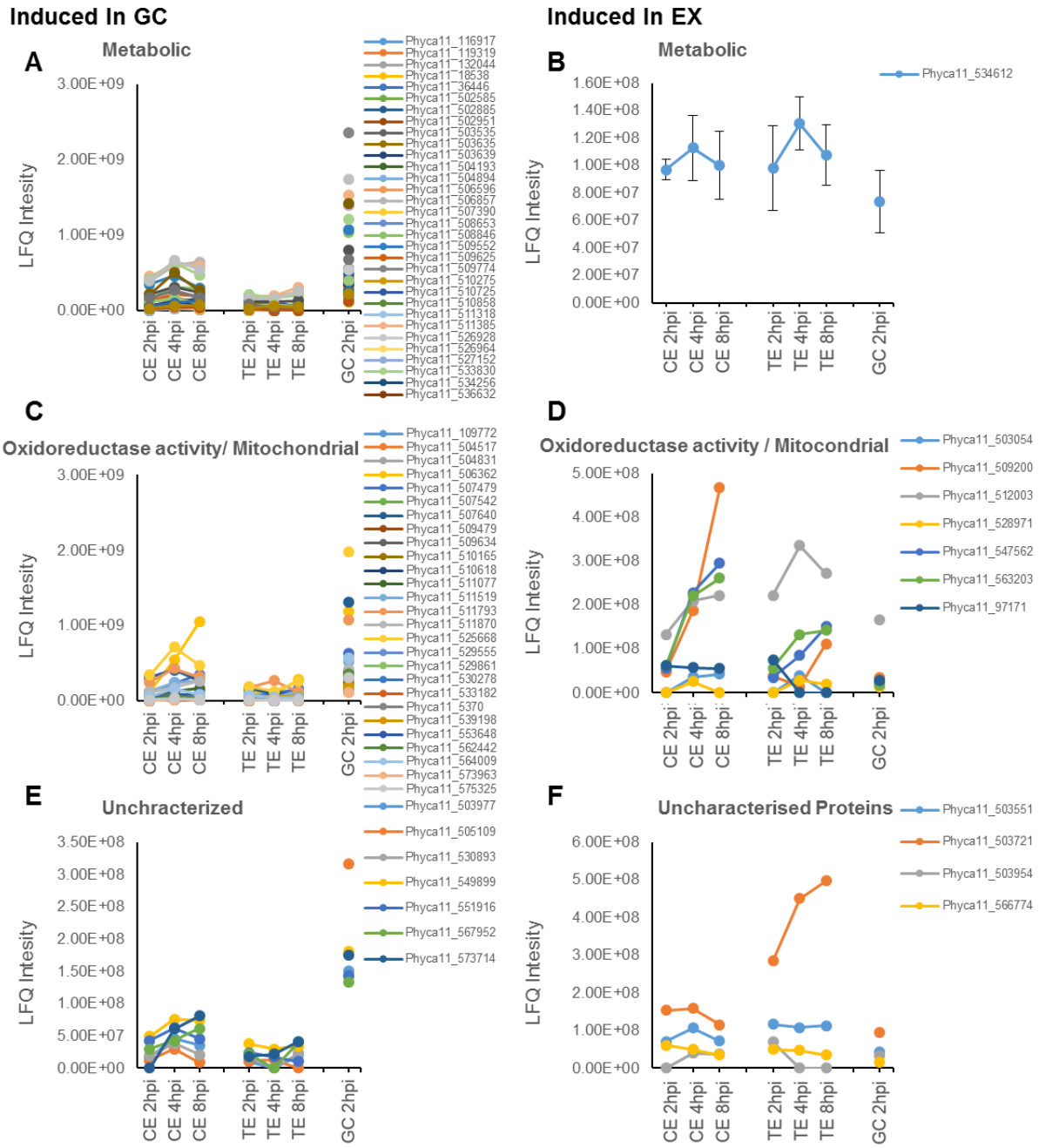


Figure 4.11: Intensity profile of Proteins.

For each protein that was manually annotated with the broad functionality category Metabolic Proteins (A,B), Proteins with Oxidoreductase activity or Mitochondrial association (C,D), and Uncharacterised Proteins (E, F), the average intensity from all replicates for each condition (CE 2, 4, 8 hpi, TE 2, 4, 8 hpi, and GC 2 hpi) has been plotted. The protein from these functionality categories that were induced in GC (A,C,E) and induced in Extract (B,D, F) are shown here. If there are fewer than 4 proteins plotted on a single graph the standard deviation has been shown with error bars.

chaperone HSP found with greater abundance by extract when compared to GC, decreases in intensity as the time course of the experiment progresses (Figure 4.9 D).

Another group of proteins that were found commonly amongst those with greater abundance in GC were related to proteasomal function. 8 (5.8%) were identified with greater abundance in GC samples (Figure 4.8), and these were KOG annotated as either a 26S proteasome regulatory complex subunit or 20S proteasome regulatory subunits (Supplementary Table 3). Contrastingly, only 1 protein with a possible proteasomal related KOG annotation was found with greater abundance in host extracts (Figure 4.8)(Figure 4.9 F). But, of those 8 proteins with increased intensity in GC, theintensity pattern is as expected; all have high abundance in GC compared to the two host extracts, and at a gross level all of these proteins seem to also have a small increase in intensity in CE compared to TE (Figure 4.9 E).

Proteins with a KOG annotation with a function in cellular structure seem to be found in similar numbers of proteins with greater abundance in both GC and extracts (Figure 4.8). However, the three (8.8%) proteins found with increased intensity in extract were KOG annotated as either beta tubulin or alpha tubulin, whereas the 11 (7.9%) proteins with increased intensity in GC (Figure 4.10 A) have a variety of annotated functions (Supplementary Table 3). Interestingly the 3 tubulin proteins all have a similar intensity pattern, decreasing in abundance in both host extracts as the time course of the experiment progresses (Figure 4.10 B).

One group of proteins that are found to a greater extent with greater abundance in extracts, compared to the number found with greater abundance in GC are those with a potential function in signalling and regulation. With 4 (2.9%) with increased intensity in GC and 7 (20.6%) with increased intensity in response to host extracts (Figure 4.8).

They vary in the KOG annotated functions, however, 3 of the 7 with greater abundance in extracts have a KOG annotation as part of the RAB GTPase superfamily (Supplementary Table 3). Intensity patterns of these 7 proteins vary, however they seem in to decrease in intensity in both extracts as the time course of the experiment progresses (Figure 4.11 C-D).

Contrasting to those there were another group of proteins that are found uniquely with greater abundance in GC. These were the 5 (3.6%) Chaperone Complex components (Figure 4.8). Although they show higher intensity in GC samples they also all show increased intensity in the extract samples as the experiment progressed, especially in CE samples (Figure 4.10 E).

The last two groups of proteins to cover were the largest two groups of those proteins that were found with greater abundance in GC, those with a metabolic function of which 149 (33.1%) were found with increased intensity in GC, and those with oxidoreductase function or mitochondrial relationship of which 27 (19.4%) were found with increased intensity in GC. Interestingly, whilst there was only 1 protein with metabolic function with greater abundance in extracts, oxidoreductase proteins were found at in similar amounts with greater abundance in extracts as GC, with 7 (20.6%) found with increased intensity in extracts (Figure 4.8). For these few proteins, graphing the intensity show that the single protein with metabolic activity peaks at 4 hpi in both extracts; the 7 oxidoreductase proteins have a variety of intensity patterns, but many show progressive increases in intensity in both extracts as the time course of infection continues (Figure 4.11 B, D).

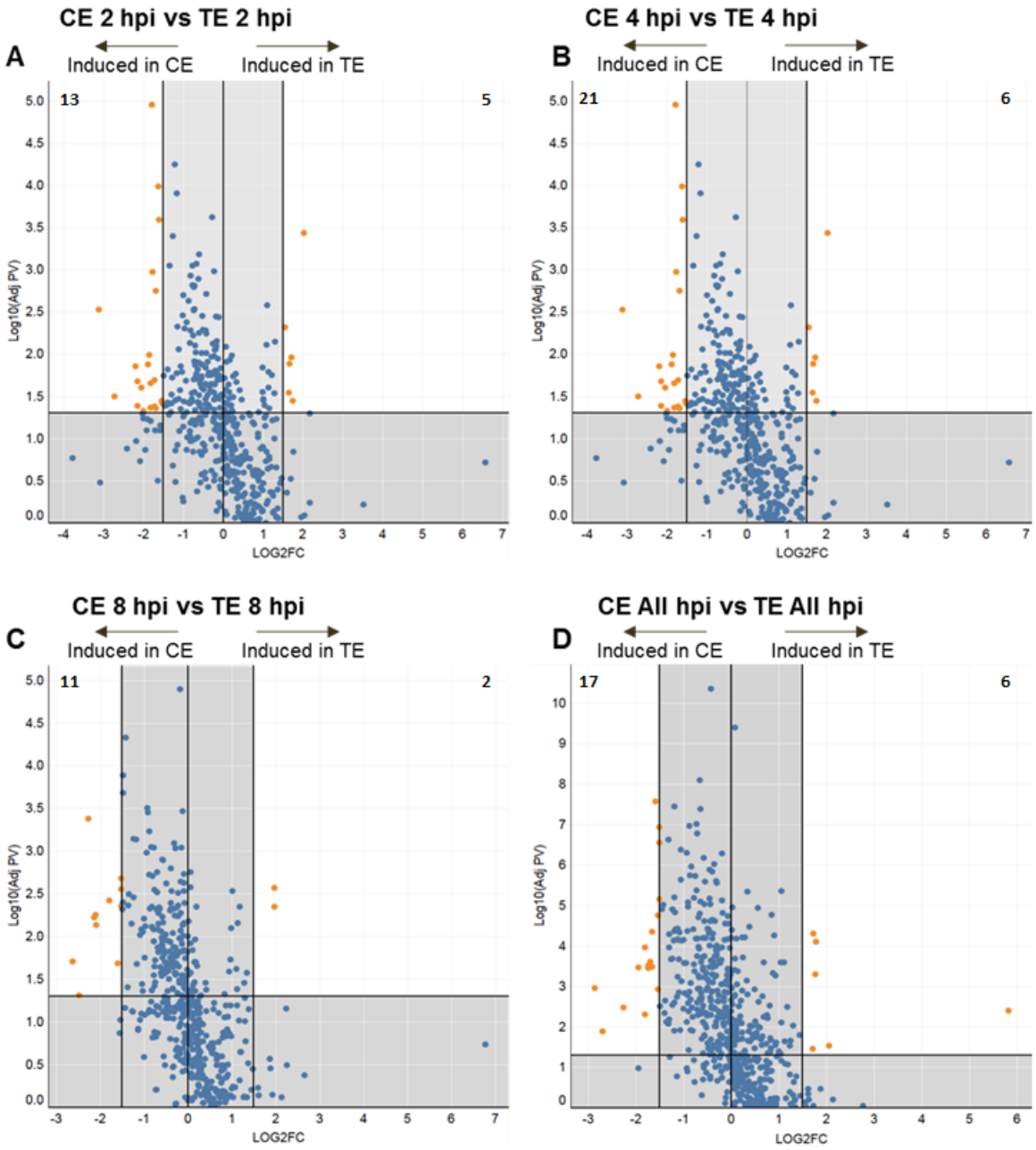


Figure 4.12: Volcano Plots showing proteins intensity for comparisons between Host extracts.

For four of the comparisons (CE vs TE at 2phi, 4hpi, 8hpi, and at all the time points pooled) the LOG2 fold change (FC) and the adjusted P value (from a two tail t-test with bonferroni correction) for each protein in found in at least 2 replicate form the two compared conditions was calculated. The LOG10 p value and the LOG2FC for each protein for each comparison have been plotted. Those proteins that have been shifted to the right are comparably induced in TE and those proteins that have shifted to the left are induced in TE. Those significantly differentially induced proteins i.e. with a Log2FC ≤-1.5 or ≥1.5, and a -log10 P value ≥ 1.3 have been shown in orange, and the point where a protein is significant has been indicated with a black line. Protein in the greyed area represented by a blue dot are not significantly differentially expressed (A-D). The number of significantly induced proteins in each subset are shown at the top of each plot.

4.2.5 Host to Host Comparison: differential protein abundance

It was also of interest to examine the differences in proteome between the two hosts. For this, 4 comparisons were made CE v TE at 2 hpi, CE vs TE at 4 hpi, CE v TE at 8 hpi, and TE v CE where all time points were pooled (All hpi). To explore this data volcano plots were generated for each comparison (Figure 4.12). For all time points, there is a defined shift to the left representing a greater number of proteins with greater abundance in CE compared to TE (Figure 4.12A-D). As before for each of these comparisons, those

proteins that were significantly differentially abundant were isolated (Log2FC ≤-1.5 or ≥1.5, and -log10 P-value ≥ 1.3). As shown by the volcano plots proteins with increased intensity in CE are in the majority of significantly differentially abundant proteins. There are 13, 21, 11 and 17 proteins which have a significantly greater abundance in samples inoculated with CE compared to TE at 2 hpi, 4 hpi, 8 hpi, and all hpi respectively (Figure 4.13 A). This is contrasted with the 5, 6, 2 and 6 proteins, at 2 hpi, 4 hpi, 8 hpi and all hpi respectively, which have a greater abundance in TE than in CE (Figure 4.13 A).

By these comparisons, the extracts were seen to have a host-dependent effect on protein abundance in *P. capsici*. In the case of proteins with a higher intensity in TE, only 11 unique proteins were identified to be significant in at least one of the time points examined. Looking at how these 11 proteins overlap at each time point shows that only two protein were found to be differentially abundant across all time points in TE (Figure 4.13 B), they were Phyca11_508140 and Phyca11_508411. A single Protein was identified differentially abundant in TE at 2 hpi and 4 hpi, as well as the pooled samples,

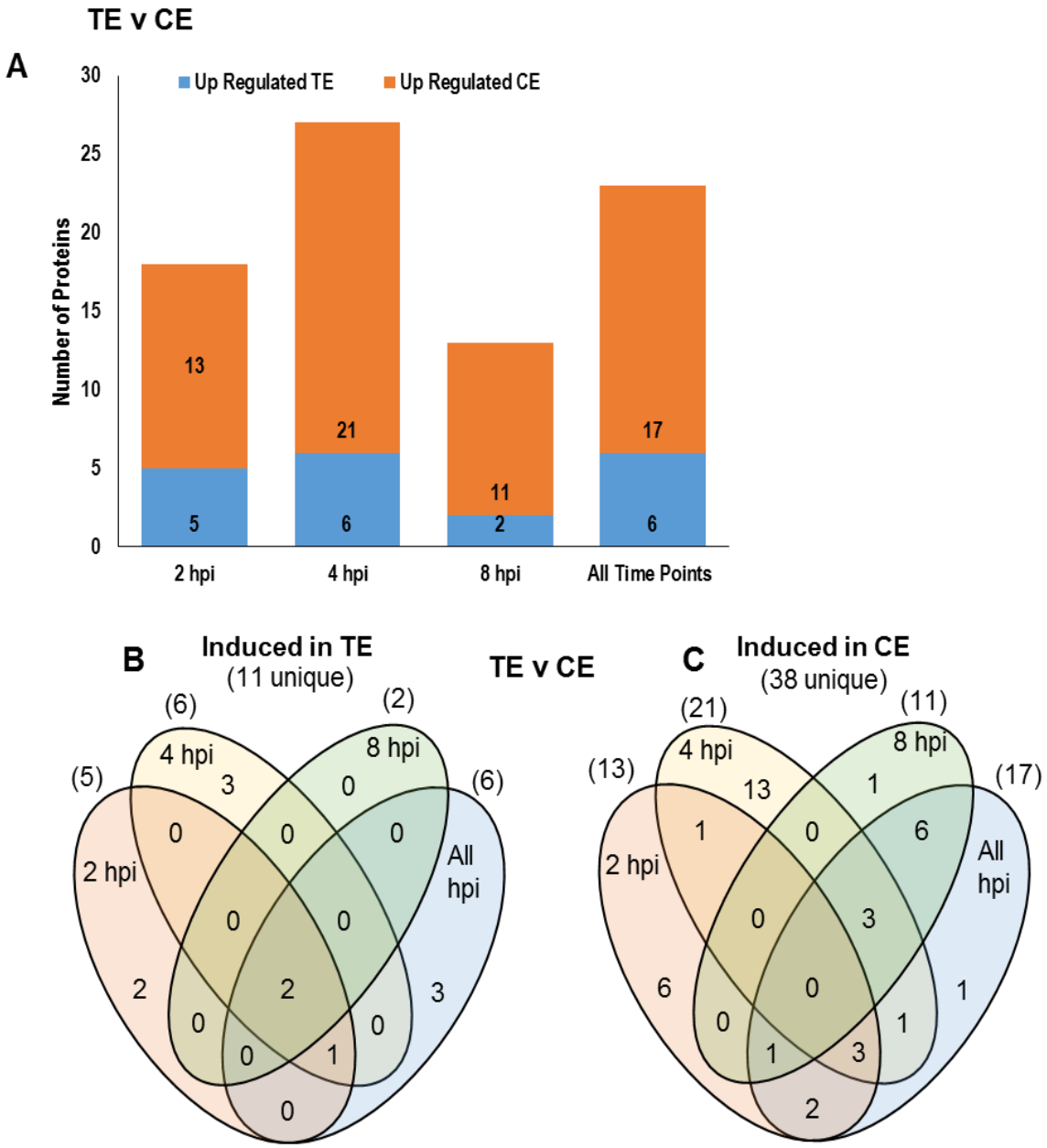


Figure 4.13: Overlap of Host – Host comparisons and Extracts GC comparisons.

The number of significantly differentially expressed proteins in each comparison have between CE and TE (those being at 2hpi, 4hpi, 8hpi and all time point pooled) have been shown with those up-regulated in CE shown in blue in the bottom portion of the bar, and those up-regulated in TE shown in orange in the top portion of the bar (A). The overlap in the proteins that were shown to be significantly induced in TE at each time point compared to CE (B) and the overlap in the proteins that were shown to be significantly induced in TE at each time point compared to CE (C) is displayed. The total number of proteins in each comparison is shown in brackets by the circle and the total number of unique proteins in the three comparisons together is shown in brackets under the title.

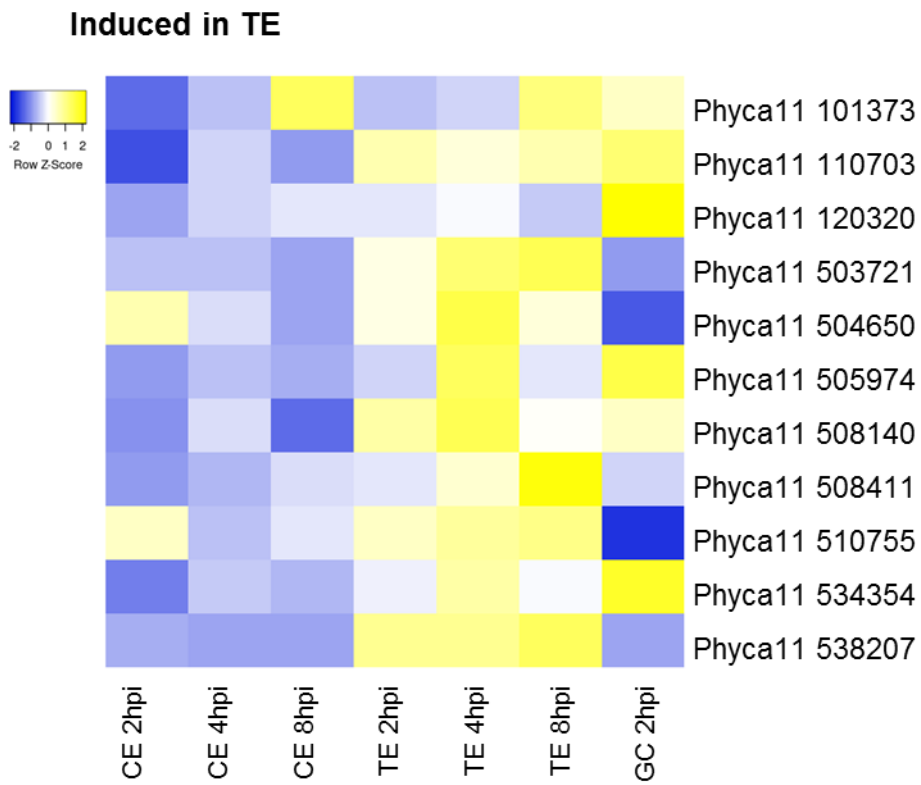


Figure 4.14: Heat map of 11 proteins with greater abundance in TE compared to CE.

Intesity is shown by adjusted z-score. Down regulation is shown in blue and up regulation is shown in yellow

Phyca11_534354. However, all other proteins with a higher abundance in TE compared to CE were only found at a single time point; 2 with greater abundance at 2hpi, 3 at 4hpi, and additional 3 proteins when the time points were pooled (Figure 4.13 B).

More proteins were identified to have greater abundance in CE than TE. There were 38 unique proteins found with significantly higher abundance in CE compared to TE. Contrasting with the proteins with greater abundance in TE, none of these were identified to have greater abundance at all time points, although many were found in just two or three of the time points. 4 hpi has the most unique proteins with greater

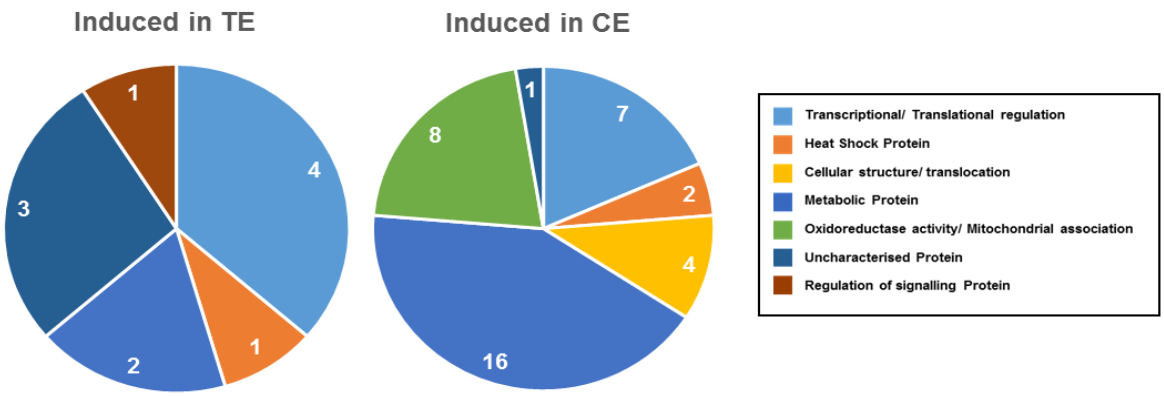


Figure 4.15: Distribution of broad functionality categories in proteins DE between hosts.

Using GO terms and KOG annotation each protein was manually assigned to a broad functionality group that was enriched in the differentially regulated proteins. These consist of 39 proteins up-regulated in CE and 11 proteins up-regulated in TE. From the top of both charts moving clockwise those categories are Transcriptional/ Translational regulation, Heat Shock Proteins, Proteins involved in Cellular structure or translocation, Metabolic Proteins, Proteins with Oxidoreductase activity or Mitochondrial association, Uncharacterised Proteins, Regulation or signalling Proteins. Note that of those proteins induced in CE none are Regulation or signalling Proteins. And of those proteins induced in TE none are Proteins involved in Cellular structure or translocation, or Proteins with Oxidoreductase activity or Mitochondrial association.

abundance at that time, having 13 compared to the 6 at 2hpi and 1 at 8 hpi (Figure 4.13 C).

Heat maps can show these same patterns as well. The 11 proteins that have greater abundance in TE have a variety of intensity patterns, but for the most part, it can be seen that for all protein, there is a higher intensity in TE lower intensity in CE, and a variety of intensity patterns in GC (Figure 4.14). However, the 38 proteins with greater abundance in CE have a consistent intensity pattern. Notably that although intensity is higher in CE, especially at 4 hpi and 8 hpi, in all but two of the proteins, there is a greater intensity in GC than in CE (Figure 4.16).

4.2.6 Host to Host comparison: Protein Functionality and intensity profile

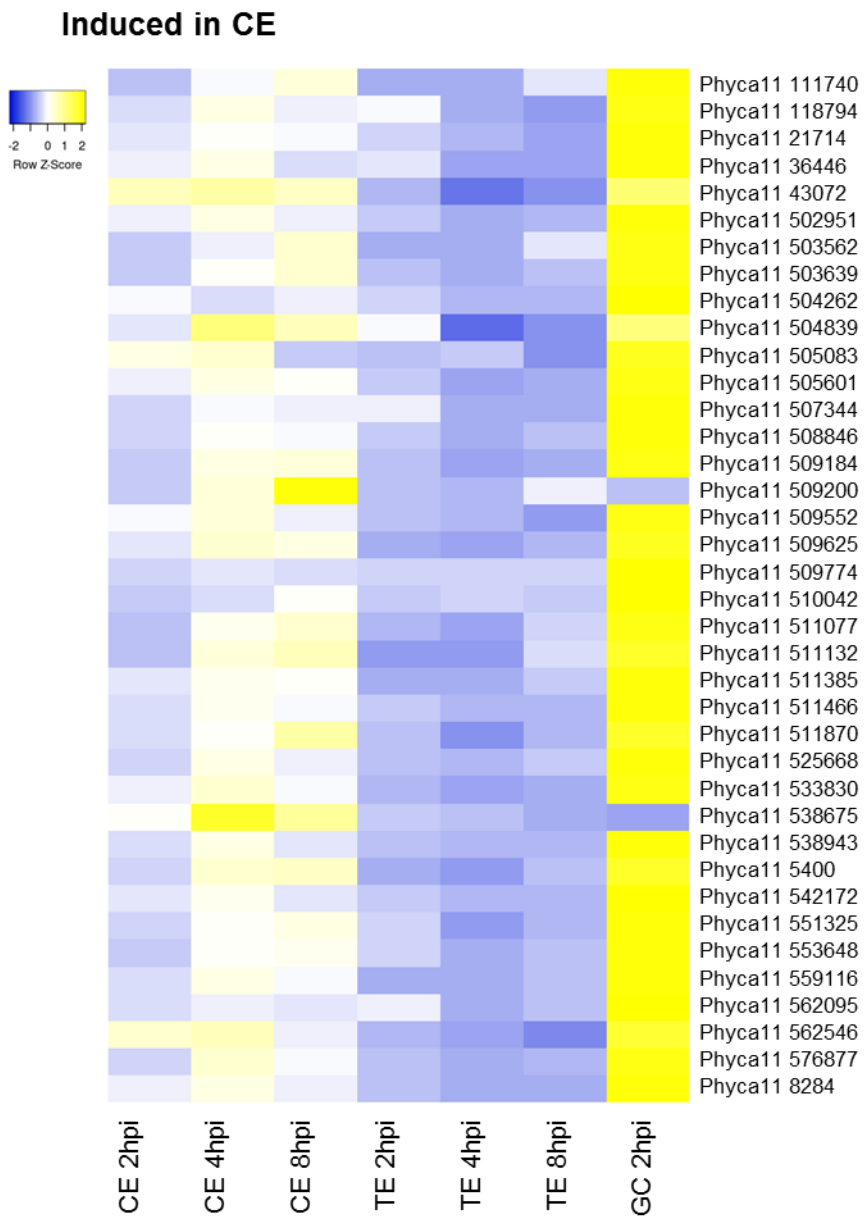


Figure 4.16: Heat map of 38 proteins with greater abundance in CE compared to TE.

Intesity is shown by adjusted z-score. Down regulation is shown in blue and up regulation is shown in yellow.

As with the previous suite of proteins, in order to explore the potential roles of these proteins, Biological process GO terms, and EuKaryotic Orthologous Groups or KOG annotations were found for each. The proteins were subdivided into the same 9 different categories. The categories were, 1) transcriptional / translation regulation, 2) Heat Shock Proteins, 3) Proteasomal, 4) Cellular structure / translocation, 5)

Regulation/ signalling, 6) Chaperonin complex component, 7) Metabolic, 8) oxidoreductase activity/ mitochondrial, 9) Uncharacterised (Figure 4.15). Challenging zoospores with both TE and CE increase the abundance of a number of proteins involved in transcription and translation, and heat shock proteins and although the proteins with increased abundance in both groups also contain metabolic proteins, only 18.2% of the proteins with increased abundance in TE had predicted metabolic activity, compared to the 42.1% of proteins with increased abundance in CE. The rest of the subcategories of proteins, bar uncharacterised proteins, only appear to have increase abundance in either TE on CE but not both. Cellular structure proteins and oxidoreductase activity proteins, only appear with increased abundance in CE, and proteins with a potential role in signalling and regulation are found only with increased abundance in TE (Figure 4.15).

The LFQ intensity of each of the host extract depended differentially abundant proteins, both those that were found with increased abundance in TE and those found with increased abundance in CE has been plotted. The intensity pattern may give clues to the mechanisms and importance of the proteins of interest.

11 proteins had a higher intensity when the *P. capsici* was challenged with TE. In all these proteins the intensity pattern differs, however, 3 generalisations can be made. For Phyca11_504650, Phyca11_534354, and Phyca11_505974, a peak of high intensity and TE 4 hpi seems to be the main contributor to their differential abundance (Figure 4.17 B, C, D). For others (Phyca11_510755, Phyca11_11073, Phyca11_508140, and Phyca11_120320) intensity remain relatively level across the three-time points, but is consistently higher than that of abundance in CE (Figure 4.17 A, B, D, E). And lastly, there are those proteins for which the intensity continue to rise

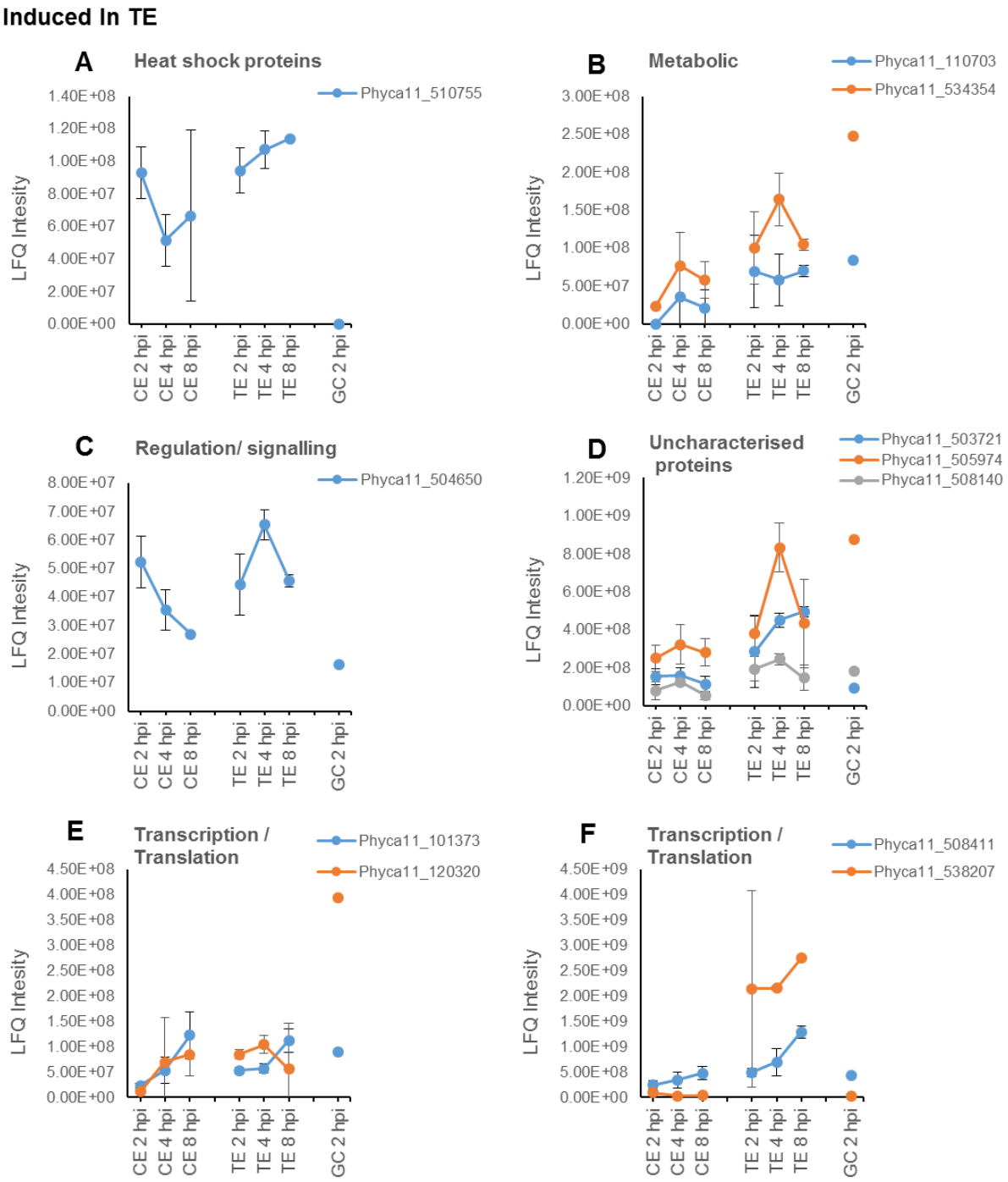


Figure 4.17: Intensity patterns and Functionality of Proteins Up regulated in TE compared to CE.

For each protein that was manually annotated with the broad functionality category Heat Shock Proteins (A), Metabolic Proteins (B), Regulation of signalling Proteins (C), Uncharacterised Proteins (D), Transcriptional/ Translational regulation (E,F), the average intensity from all replicates for each condition (CE 2, 4, 8 hpi, TE 2, 4, 8 hpi, and GC 2 hpi) has been plotted. The protein from these functionality categories that were induced in TE are shown here. If there are fewer than 4 proteins plotted on a single graph the standard deviation has been shown with error bars.

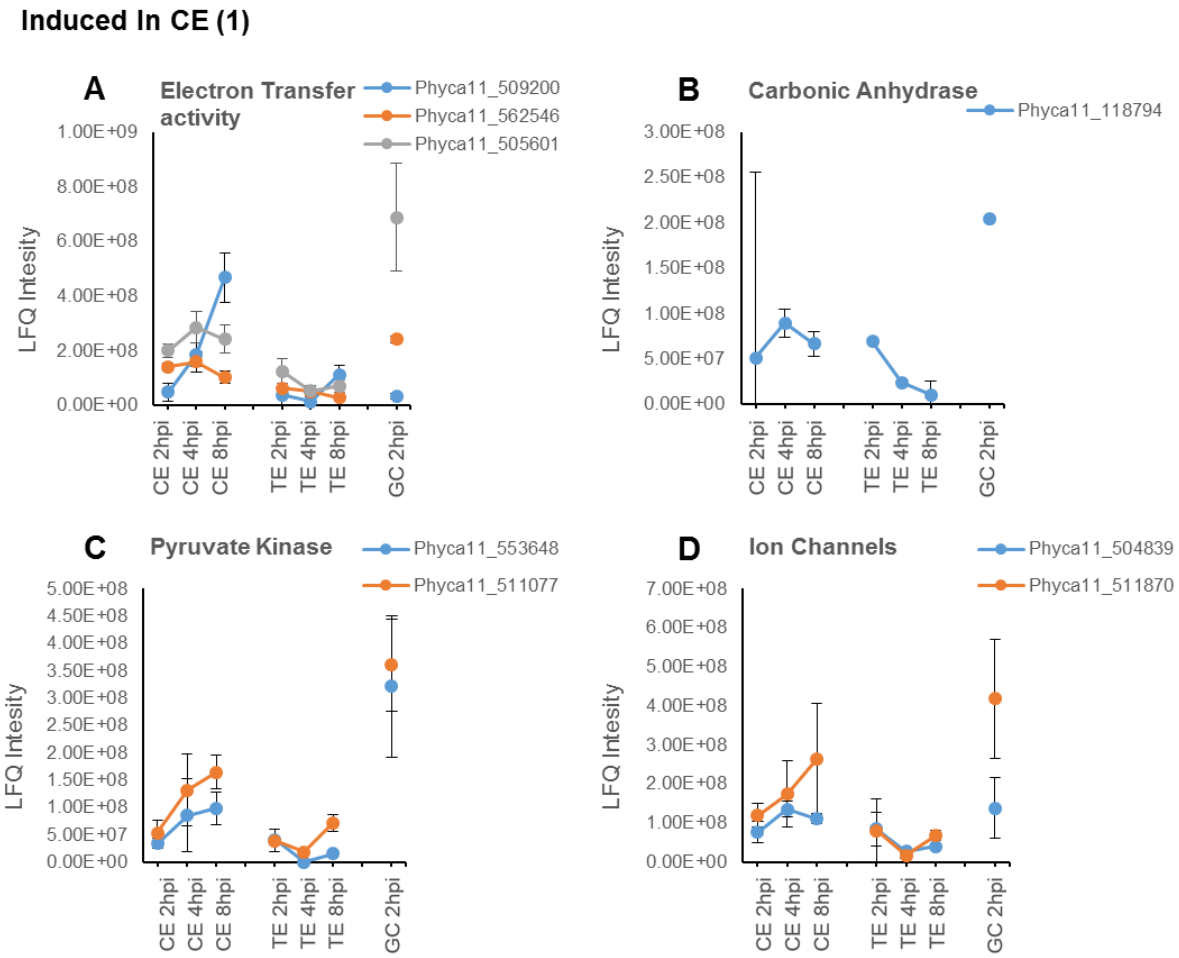


Figure 4.18: Intensity patterns and Functionality of Proteins Up regulated in CE compared to TE.

For each protein that was manually annotated with the broad functionality category Oxidoreductase activity or Mitochondrial association, the average intensity from all replicates for each condition (CE 2, 4, 8 hpi, TE 2, 4, 8 hpi, and GC 2 hpi) has been plotted. The Proteins have been further broken sown into subcategories. The protein from this functionality categories that were induced in TE are shown here. If there are fewer than 4 proteins plotted on a single graph the standard deviation has been shown with error bars.

over the course of the experiment, Phyca11_503721, Phyca11_101373, Phyca11_508411, and Phyca11_538207 (Figure 4.17 D, E, F). The last of these Phyca11_538207 stand out by having a dramatic and consistent increase intensity in TE compared to its intensity in the other host extract and its

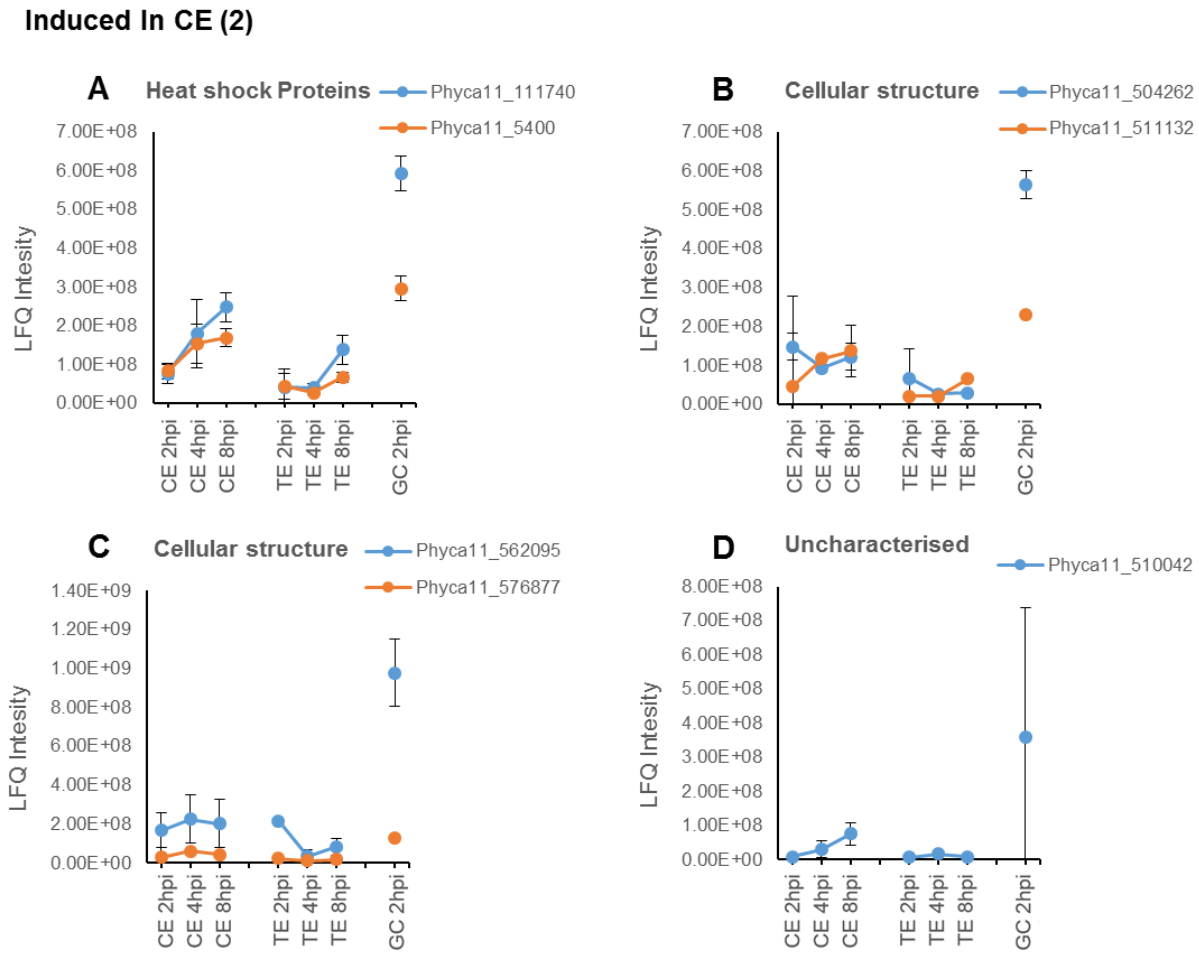


Figure 4.19: Intensity patterns and Functionality of Proteins Up regulated in CE compared to TE.

For each protein that was manually annotated with the broad functionality category Heat Shock Proteins (A), Proteins involved in Cellular structure or translocation (B,C), and Uncharacterised Proteins (D), the average intensity from all replicates for each condition (CE 2, 4, 8 hpi, TE 2, 4, 8 hpi, and GC 2 hpi) has been plotted. The protein from these functionality categories that were induced in CE are shown here. If there are fewer than 4 proteins plotted on a single graph the standard deviation has been shown with error bars.

intensity in GC (Figure 4.17 F). There seems to be little correlation between the intensity pattern and the subcategory the proteins were placed in due to their predicted function, except for the fact that the three of the four proteins with a potential role in transcription and translation both show progressive increases in intensity over the time course of the experiment (Figure 4.17 F). Those being Phyca11_101373,

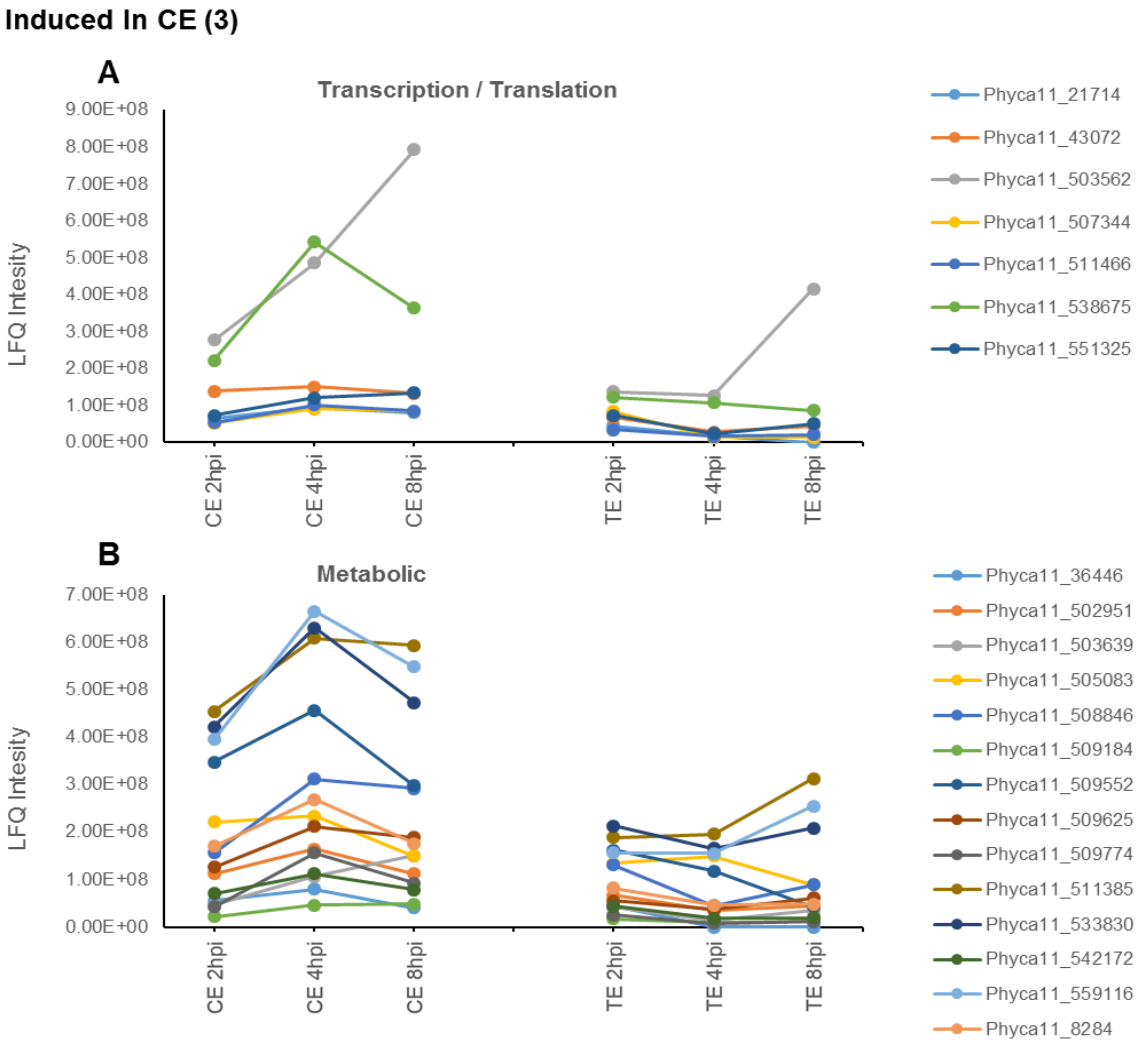


Figure 4.20: Intensity patterns and Functionality of Proteins Up regulated in CE compared to TE.

For each protein that was manually annotated with the broad functionality category Transcriptional/ Translational regulation (A), Metabolic Proteins, (B), the average intensity from all replicates for each condition (CE 2, 4, 8 hpi, TE 2, 4, 8 hpi, and GC 2 hpi) has been plotted. The protein from these functionality categories that were induced in CE are shown here.

Phyca11_508411, and Phyca11_538207 KOG labelled as a 40S ribosomal protein S2, a 26S proteasome regulatory complex subunit, and a DEAH-box RNA helicase respectively (Supplementary Table 4). For the 38 proteins that have a higher intensity when *P. capsici* was challenged with CE, 8 were found to have oxidoreductase activity or to be associated with the mitochondria. For graphing the intensity patterns of those proteins involved in oxidoreductase activity/ mitochondrial, have been subdivided

again, into ion channels, pyruvate kinases, electron-transferring activity and a carbonate dehydratase (Supplementary Table 4). In addition to these 8 proteins, 2 heat shock protein were identified, 4 proteins involved in cellular structures and 1 uncharacterised protein's (Figure 4.15)(Supplementary Table 4). The two larger groups are the 7 proteins with potential roles in transcription and translation, and the 14 proteins with potential metabolic activity. For those 8 proteins with a potential role in oxidative stress, oxidoreductase processes and/or mitochondrial processes, for the most part had a similar intensity pattern. The proteins all have higher intensity when challenged with CE, often increasing in abundance as the time course progresses, whereas the abundancy found in TE either reduce or remain relatively level (Figure 4.18). A similar pattern of increasing intensity in by CE as the time course progresses can also be noted in the two heat shock proteins, Phyca11_511132 a potential nucleosome assembly protein, and the uncharacterised protein Phyca11_510042 (Figure 4.19 A, B, D). For the other three proteins with a predicted role in cellular structure, their intensity is relatively stable across the time points, with consistently higher intensity in by CE (Figure 4.19 B, C). Similar patterns to those described above can be seen in the 7 proteins with potential roles in transcription and translation, and the 14 proteins with a potential metabolic activity (Figure 4.20). It should be noted that for all these proteins with increased intensity in CE compared to TE, with the exceptions of Phyca11_509200 (Figure 4.18) and Phyca11_538675 (Figure 4.20) there is a higher intensity in GC than in CE or TE (Figure 4.15, 17-19)

4.2.7 Comparison of the Proteomics data set with RNA-SEQ and Microarray

To both validate and examine the selectivity and sensitivity of both this proteomic data and the RNA sequencing transcriptomic data from chapter 4, comparisons of

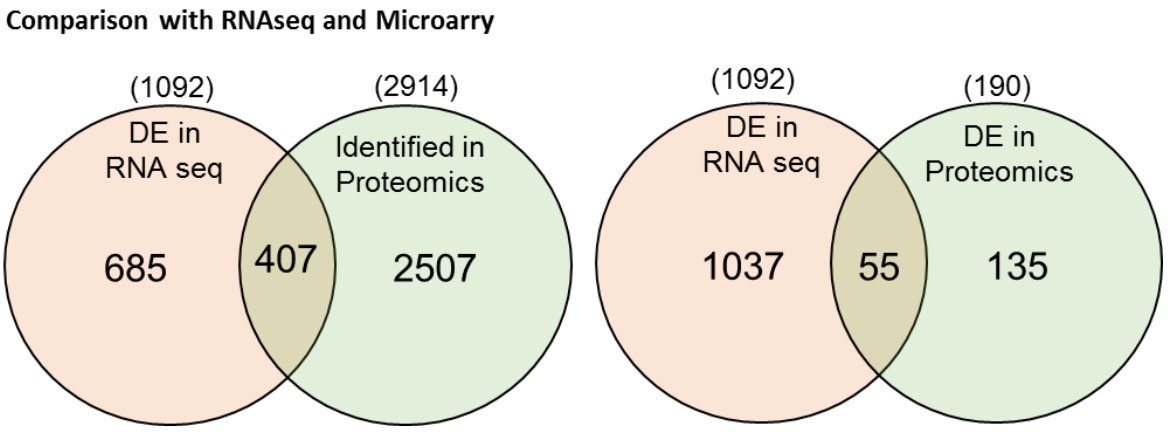


Figure 4.21: Proteomic comparison with RNA sequencing and Microarray.

The overlap of those all proteins identified at in at least on condition, and the differentially expressed (DE) genes from the RNA sequencing (Chapter 4) (A). And the over lap of all differentially abundant proteins from all comparisons and the differentially expressed genes from the RNA sequencing (Chapter 4) (B). The total number of proteins/genes in each group is shown in brackets above the Venn diagrams.

differentially abundant proteins sets and the differentially expressed gene sets were made. Immediate comparisons showed that of all the genes identified and found to be DE in from the RNA sequencing experiments of Chapter 4, of those 1092 genes, 407 were identified in the Proteomic set. Of those 407 proteins that had corresponding differentially expression in the RNA sequencing, 55 were found to also be differentially abundant in the proteomic data set (Figure 4.21). 42 of the 55 genes/proteins that were differentially expressed in both RNA sequencing and proteomics were from those genes that were according to the RNA sequencing induced in extracts, leaving 13 that were found with increased intensity in GC. 3 of these were also DE in the RNA sequencing when comparing between CE and TE (Table 4.1).

When comparing those DE genes identified by the RNA-SEQ to the differentially abundant proteins in the proteomics, first looking at the largest group; the 42 genes induced in extract compared to GC, it was found that 30 of those were actually differentially induced in GC compared to higher abundance in extract in the proteomic

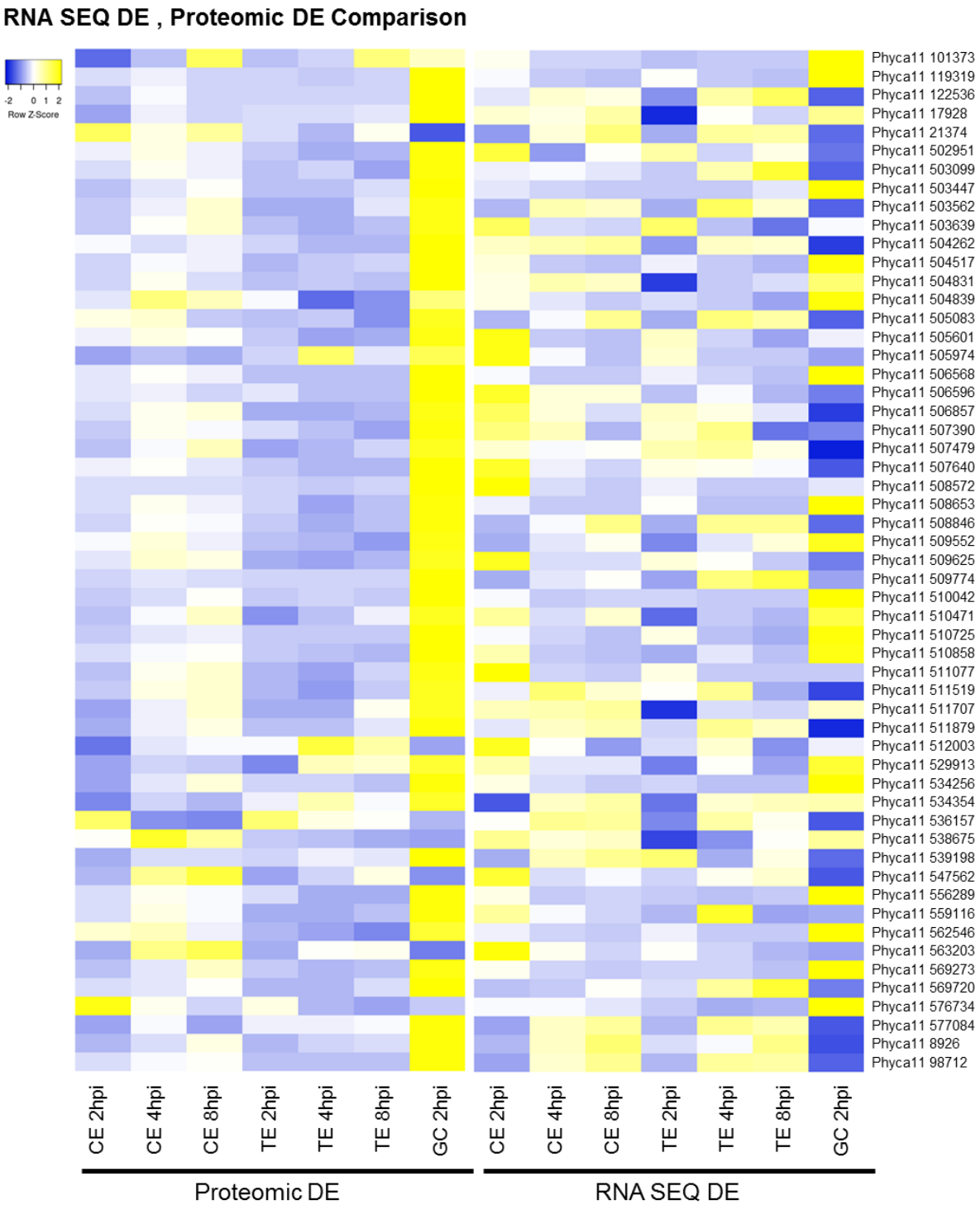


Figure 4.22: Heat map of 55 DE expressed genes from RNA sequencing that can also be found DE in Proteomics.

Expression is shown by adjusted z-score. Down regulation is shown in blue and up regulation is shown in yellow.

data, and only 6 were found to match the RNA-SEQ by also being with higher abundance in GC. 14 of the genes were also found to be DE in the host-to-host

Table 4.1: **RNA seq vs Proteomics DE genes/proteins overlap.**

The 55 Proteins/genes that were both differentially expressed in the proteomics and the RNA sequencing. The conditions in which these proteins and genes are differentially expressed (DE) is shown with a tick in the top table. In the summary table below the total number of DE genes or proteins induced in either TE, CE, Extract or GC is shown from both the RNA SEQ and the proteomics. Furthermore the number of the 55 genes/proteins that were found in DE in both RNA SEQ and proteomics that are found overlapping in at least two of the groups of induced gene when comparing the RNA SEQ to the Proteomics

	RNA Sequencing				Proteomics			
Phyca11 #	UP in CE	UP in TE	UP in GC	UP in EX	UP in CE	UP in TE	UP in GC	UP in EX
502951				✓	✓		✓	
503562				✓	✓		✓	
503639				✓	✓		✓	
504262				✓	✓		✓	
508846			✓		✓		✓	
509552			✓		✓		✓	
509625				✓	✓		✓	
509774			✓		✓		✓	
511077				✓	✓		✓	
534354	✓			✓		✓	✓	
538675			✓		✓			✓
559116				✓	✓		✓	
101373				✓		✓		
119319			✓				✓	
122536				✓			✓	
17928	✓		✓				✓	
21374		✓		✓				✓
503099				✓			✓	
503447				✓			✓	
504517				✓			✓	
504831			✓				✓	
504839				✓	✓			
505083				✓	✓			
505601				✓	✓			
505974				✓		✓		
506568				✓			✓	
506596			✓				✓	
506857				✓			✓	
507390			✓				✓	
507479				✓			✓	
507640			✓				✓	
508572				✓			✓	
508653			✓				✓	
510042			✓		✓			
510471				✓			✓	
510725			✓				✓	
510858				✓			✓	
511519				✓			✓	
511707				✓			✓	
511879				✓			✓	
512003				✓				✓
529913				✓			✓	
534256				✓			✓	
536157				✓				✓
539198				✓			✓	
547562				✓				✓
556289				✓			✓	
562546				✓	✓			
563203				✓				✓
569273				✓			✓	
569720				✓			✓	
576734				✓				✓
577084				✓			✓	
8926				✓			✓	
98712				✓			✓	

		RNA TE vs CE		RNA EX vs GC	
		Induced in CE	Induced in TE	Induced in EX	Induced in GC
Proteomics	Total	110	29	371	665
Induced in TE	11	1	0	3	0
Induced in CE	38	0	0	11	5
Induced in EX	34	0	1	6	1
Induced in GC	139	2	0	30	11

comparison in the proteomic data set, 3 were induced in TE and 11 induced in CE. The comparison improves, however, when examining the 13 genes with higher expression in GC compared to extracts. The proteomics found 11 of the corresponding proteins also had higher intensity in GC, whilst only 1 of the proteins was found with higher intensity in the extract. The final gene could only be found in the proteomics data with greater abundance in CE (Table 4.1) (Figure 4.22).

4.3 Methodology

4.3.1 P. capsici and Host Plant growth and Extract Inoculation Assay

Commercially available tomato plants (*Solanum Lycopersicum* cv. Moneymaker) and cucumber plants (*Cucumis sativus* cv. Venlo Pickling) susceptible to *P. capsici* were grown with a 16 hour light cycle and maintained at 22-25°C. 3 – 5-week old tomato and cucumber leaf and stem material were used to produce extract for inoculation assays. *P. capsici* strain LT1534 was grown on V8 agar (10% v/v V8 vegetable juice, 1 g/L calcium carbonate, 300 mg/L β-sitosterol, 15% w/v Agar).

Zoospores were collected. *P. capsici* LT1534 was grown on V8 agar in the dark at 25°C sealed with parafilm. This was followed by 2 days of growth in the light at 22°C without parafilm to induce sporangia formation. Zoospores were collected by flooding the 150 mm plate with 30 mL of sterile distilled water at room temperature, then the mycelia growth was agitated with a sterile plate spreader and everything was transferred to a second *P. capsici* inoculated V8 agar plate. Continuing in this fashion, the 30 mL of water was used to flood and agitate four 150 mm plates recovered into a 50 mL falcon tube. The sporangia suspension was left for 30-45 minutes with the lid

removed on a lightbox to induce zoospore release. The suspension was then filtered through one layer of Miracloth to remove mycelia and any agar chunks in the suspension, however both full and empty sporangia are retained, this process yielded approximately 25 mL of ≥ 1 x 10^6 spores mL^{-1} zoospore suspension. The zoospore suspension was then diluted to 1 x 10^6 spores mL^{-1}.

For the Extract inoculation experiments, plant leaf and stem tissues were collected as described. To produce extracts, 20 g of plant material was flash-frozen in liquid nitrogen immediately after collection. This was then ground up using a pestle and mortar and suspended in 100 mL of sterile distilled water. This was then filtered through one layer of Miracloth and topped up to 100 mL with sterile H_2O. 5 mL of 1 x 10^6 spores mL^{-1} zoospore suspension and 5 mL of either broth, extract or distilled water were mixed in a 15 mL falcon tube. This inoculated broth, extract or water was then shaken vigorously to induce zoospore germination. The 15 mL falcon tube were stored at 22°C in the light, and placed at an angle on its side with no lid. For protein extraction, the liquid was discarded carefully, leaving germinated spores adhering to the side of the tube. It was flash-frozen in liquid nitrogen and kept at -80°C until extraction.

4.3.2 Protein extraction and Purification

Samples were removed from the -80°C freezer and placed on ice. Several extractions and purification methodologies were trialled to see which would give the best concentration and integrity and visualised by electrophoresis of total protein extraction. All extractions were conducted using GTEN buffer as a base, GTEN consists of 10% glycerol, 25 mM Tris pH 7.5, 1 mM EDTA, and 150 mM NaCl. Using GTEN as a base two extraction buffers were trialled GTEN-T and GTEN-DM, GTEN-T consists of GTEN, 2% w/v PVPP, 10 mM DTT, and 0.1% v/v Tween 20. Whereas GTEN-DM

consist of GTEN, 2% w/v PVPP, 10 mM DTT, and 1% v/v detergent mix. Detergent mix consists of 25% v/v Brij L23, 25% v/v Triton X-100, 25% v/v IGEPAL CA-630, 25% v/v TWEEN 20. 5ml of extraction buffer (GTEN-T or GTEN-DM) was added to the 15ml falcon tube containing the sample, vortexed for 1 minute and then Spun at 12,000 xg at 4°C for 10 minutes, collecting the supernatant following centrifugation. Or, the 5ml was vortexed for 1 minute, and then decanted into 3 2 mL Eppendorfs containing steel balls. This was then placed in a TissueLyser II (Qiagen) and lysed at 30 Hz for 1 minute. Samples were then centrifuged at 12,000 xg at 4°C for 10 minutes and the supernatant collected.

Supernatant for both extraction using GTEN-T or GTEN-DM, and both using vortexing and the TissueLyser II to aid lysis were both loaded on a gel for electrophoresis or underwent methanol purification. In a 2 mL Eppendorf, to 200 µL of sample 800 µL of Methanol was added, and then vortexed to mix. 200 µL of Chloroform was added and again vortexed to mix. 600 µL of Milli-Q water was added and again vortexed to mix. The mixture was centrifuged for 2 min at 14,000 xg. The top aqueous layer was removed and discarded, followed by the addition of a further 800 µL of Methanol and the mixture inverted several times to mix. The mixture was centrifuged for 3 min at 14,000 xg and the supernatant was pipetted off. The pellet was allowed to air dry for 5 minutes and then suspended in Tris-SDS (pH 7.5) (25 mM Tris-Cl, and 0.1% SDS)

For electrophoresis, 15 µL of the sample was added to 5 µL of x4 sample buffer containing 0.2M Tris-HCl (pH 6.8), 8% w/v SDS, 6mM Bromophenol blue, and 40% v/v Glycerol. 0.5 µL 1M DTT was added to each 20 µL sample and the samples were boiled at 98°C for 5 minutes. Samples were then loaded onto a 4–15% Mini-PROTEAN® TGX™ Precast Protein Gels (BioRad) and run per manufactures instructions. The gel was then fixed with 50% v/v Methanol, 10% v/v Glacial acetic

acid, and 40% v/v water overnight, then stained with the above solution containing with 0.25% Coomassie Brilliant Blue R-250 for 2 – 3 hours. And, de-stained overnight in 5% v/v Methanol 7.5% v/v Glacial acetic acid, and 87.5% v/v water.

For Proteomic analysis samples were extracted as above using GTEN-DM, lysed as above by simply vortexing, purified as above using methanol precipitation and suspended in Tris-SDS.

4.3.3 Mass Spectrometry Analysis and Data analysis

Mass spectrometry was conducted by the University of Dundee FingerPrints Proteomics Facility. S-TRAP purification was conducted on all samples. Following this a minimum of 1 µg of peptides from each sample was analysed using the Ultimate 3000 RSLCnano ultra-high-performance liquid chromatography system (Thermo Scientific) coupled with a Q Exative HF-X Mass Spectrometer (Thermo Scientific). Samples were then injected in 0.1% formic acid and trapped on an Acclaim PepMap 100 (C18, 100 µM x 2 cm). The peptides were then subsequently separated on an Easy-Spray PepMap RSLC C18 column (75 µM x 50 cm) (Thermo Scientific). The mobile phase consisted of 0.1% formic acid (solvent A) and 80% acetonitrile in 0.08% formic acid (solvent B). A constant flow rate of 0.3 µL/min was used. Following separation, the peptides samples were then transferred to the mass spectrometer via an Easy-Spray source at 50ºC with a source voltage of 1.9 kV. The top 15 most intense peaks in each MS1 scan were then taken for MS2 analysis. The level of resolution for MS2 spectra was 17,500 and the spectra were fragmented using collision-induced dissociation (CID) with a mass range of 350 to 1600 m/z.

The RAW files were then analysed using MaxQuant version 1.6.2.0 against a P. capsici database with data manipulation in Microsoft Excel 2013 to produce comparisons between samples. Maxquant analysis and sample comparisons were conducted by C. Rogers. Fixed modifications included Carbamidomethyl (C) and variable modifications included were Oxidation (M), Acetyl (N-term), Deamidation (NQ), Dioxidiation (MW) and Gln->Pyro-Glu with an error tolerance of 10ppm for FTMS and 0.06Da for ITMS. Average intensity for each condition was calculated for the 4 biological replicates, and the fold change was calculated by dividing on condition of interest with the comparison condition. For significance tests, the log2 fold changes were calculated for each comparison and a two-sample equal variance two-tailed t-test was used to generate the P-value. Significant differential abundance Log2FC 1.5 + or – from zero and -log10 P-value of 1.3 or greater (equivalent to P-value of ≤ 0.5). Significant differential abundance was only calculated if at least two of the replicates in each condition had intensity data.

4.4 Discussion

Proteomics and other omics-based studies have examined many facets of the host-pathogen interaction, and the many stages of the pathogen life cycle (Pang et al., 2017, Pang et al., 2016, Reynoso et al., 2015). However, there is a gap in these studies that remains unfilled. An omics-based approach to the examination of the infectious life-cycles stages and notably the pathogens side of the pathogen-host interaction remains elusive. This is partly because infecting pathogen is imbedded into its host plant's tissue, and is hard to isolate and study in isolation, but in combination plant material is in such excess it makes omics style experiments impractical. In an attempt to fill this gap in current research and to examine the mechanism of host

adaption of *P. capsici*, especially in the context of distinct host plants, in Chapter 3 an extract inoculation assay was explored. This methodology attempted to partially model the early stages of plant infection and host perception to allow transcriptomic and proteomic analysis of potential mechanisms and key genes/proteins in the mechanisms host adaptation. Proteomic analysis of germinating cysts and the early time points of zoospore germination, and the formation of infection machinery in a model of a plant infection, has revealed a suite of interesting proteins with potential roles in host adaption.

4.4.1 Cysts germinating in extract versus water control (Extract v GC): metabolism and respiration

Although the experiment was designed to explore the differences induced by the host, the most dramatic shift in protein abundance happened when comparing GC to extract. This pattern of a large induction event in GC compared to *P. capsici* inoculating host extract was also observed in the RNA sequencing. The Proteomics suggests a mechanism as to why this may be the case. In fact, it may be the case that cysts germinating in water are a poor control condition for the examination of host-based differences in cyst germinating a plant extract.

The two largest groups of proteins that were found with increased intensity in GC were proteins with a KOG annotation and or GO term that associated them with 1) metabolism or 2) oxidoreductase activity/ mitochondrial activity. This suggests a change or increase in respiratory activity paired with a change in the nutrients the cell has available or is using. Metabolic pathways that were frequently observed were those of amino acid synthesis, with multiple amino acid transferases and synthases, displaying remarkable *de novo* synthesis of amino acids (Supplementary Table 3). In fact, elevated amino acid biosynthesis has been observed previously in *P. infestans*

during appressorium formation suggesting that elevated amino acid biosynthesis and metabolism is required during germination (Grenville-Briggs et al., 2005). However, much like in this study cysts were allowed to germinate in the absence of plant material, so it is unclear if the addition of plant-derived nutrient in the form of plant extract alters the metabolic pathways usually seen during germination. In addition, there are other proteins that are elevated in GC and demonstrate a high metabolic activity in this life cycle stage; proteins that are involved the metabolism of high energy carbohydrates such as malate, succinate, and pyruvate demonstrating a high energy life stage. In addition, many elements of the fatty acid synthesis pathway were also observed to with increased intensity in germinating cysts, suggesting a change in cellular function or morphology and perhaps could be related to appressorium formation or again suggest a high metabolic activity.

Acetyl-CoA, acetyltransferase, Pyruvate Kinases, Aldehyde dehydrogenase, other enzymes commonly associated with the mitochondrial, and dehydrogenase enzymes with roles in various metabolic pathways were also found to be elevated in GC. This suggests that the mitochondria are playing an important role in cyst germination, potentially having a protective role against oxidative stress, although the increase in abundance of multiple amino acid dehydrogenase suggest a role in the *de novo* synthesis of amino acids. There is also a likely increase in mitochondrial respiration that pairs nicely with the metabolic activity observed. For example, The fatty acid degradation pathway and enzymes found in that pathway are of particular note, as they have been found with increased intensity in germinating cysts in *P. capsici* (Pang et al., 2017) and other *Phytophthora* species (Hosseini et al., 2015). In this study in GC, they found several enzymes known to be involved in the mitochondrial β-oxidation pathway of fatty acid degradation in addition to key enzymes of the glyoxylate cycle

that allows the synthesis of high energy carbohydrates for fatty acids such as malate. This led to the suggestion that lipid reserves are key energy stores used for cyst germination (Hosseini et al., 2015).

4.4.2 Cysts germinating in extract versus water control (Extract v GC): Protein regulation

One set of proteins that were almost uniquely found with greater abundance in GC were subunits of the proteasomal regulatory complex. Whilst one 26S proteasome regulatory complex subunit was found with greater abundance in extract samples compared to GC, 8 were found with increased intensity in GC. Interestingly though no known regulator of enzymes that play a role in Ubiquitination were found with increased intensity in GC, whereas a single Ubiquitin-protein ligase was found differentially abundant in extract. However, this could be part of a starvation/ stress phenotype, or it could be part of a natural nutrient recycling to maintain amino acid levels during a highly metabolically active stage of *P. capsici* lifecycle. Although germinating cysts, in the tip of appressorium have been shown to have high levels of autophagy (Luo et al., 2014), it was, in fact found in *P. infestans* that proteasomal proteins did not play a big role in cyst germination, at least shown by RNA sequencing (Ah-Fong et al., 2017).

In addition to protein regulation through the proteasome, heat shock proteins were found in relatively high numbers differentially abundant in both GC and extracts conditions. Heat shock proteins whilst being part of the normal cellular stress response to temperature and other stressor degrading damaged proteins, also function as molecular chaperones responsible for protein folding, assembly, and translocation.

These proteins have been observed to be key in plants response to pathogen infection (Park and Seo, 2015). Interestingly HSPs seem to be understudied in plant pathogens, especially in the context of the stress response to plant defence. Although some HSPs appear to be temperature responsive in some *Phytophthora* species (Puig et al., 2018). This project however, suggests that other (plant-derived) signals, or the absence thereof, may affect their abundance in germinating cysts. The extent with which they contribute to infection and colonisation, however, remains to be determined.

The presence of an active amino acid recycling mechanism pairs nicely with the dynamic, highly metabolically active cell, with high levels of protein synthesis and turnover that the rest of the proteomic suggests is occurring in GC. This does, however, beg the question; if germinating cyst are showing an active highly metabolic phenotype, with high respiration, and protein synthesis and turn over, what kind of phenotypic characteristic are found in cells challenged with extract and why do they differ from GC.

4.4.3 Cysts germinating in extract versus water control (Extract v GC): Transcriptional Regulation

Unlike GC there was no uniform regulation of metabolic proteins for plant extracts, although host difference in metabolic proteins was noted. Additionally, changes in dehydrogenases and mitochondrial proteins were noted, showing a similar although distinct increase in respiratory activity. However, the largest group of induced proteins in extract were those with a role in transcription and translation, and indeed after proteins involved in metabolism and mitochondrial processes the translational and transcriptional proteins are also the most abundant induced proteins in GC.

Examining those 26 proteins that have a predicted role in transcriptional and translation also revealed that half were tRNA synthases. *De novo* synthesis of tRNAs suggests that in GC, cellular processes, geared towards protein synthesis, may be activated in preparation of invasion. Although beyond their canonical role in protein synthesis, tRNAs may also have other regulatory functions that could affect *P. capsici* biology. For example, tRNAs have been shown as substrates for post-translational protein and cell membrane phospholipid modification (Guo and Schimmel, 2013). tRNA fragments have also been shown to have a role in translational regulation (Guo and Schimmel, 2013). While the exact role of tRNA's in GC remains to be elucidated, t-RNA synthesis plays an important role in GC and are not key when the cysts are introduced to extract. The other 13 proteins suggest further that GC are regulating translational processes, with two Translation initiation factor 3s, two ATP dependent RNA helicases, and a single Predicted RNA-binding protein containing PIN domain. Suggesting a close regulation of translation and protein production. It is clear that there is a shift in protein biosynthesis that differs between GC in water and *P. capsici* germinating in extract. Compare this to the focus in those proteins with putative role in transcriptional regulation with increased intensity in extract. 5 of the 10 are proteins that are part of the ribosomal complex, showing again a regulation of protein synthesis but in a different manner, along this line, a DEAH-box RNA helicase was also identified. Besides this translational regulation, two Histones proteins were found with increased intensity, suggesting there is also the regulation of gene expression happening in extracts that is not found in GC. Perhaps even representing a mechanism where histones are released from chromatin to induce expression of genes, backed up by the fact that no increase in transcript level for histon genes were foun the RNA sequencing experiments (Chapter 3) (Meile et al., 2020).

4.4.4 Regulatory and structural Proteins uniquely induced in Extract

Interestingly one category of proteins that were not found with increased intensity in GC but was found numerous times in extracts were proteins that were subcategorised as regulatory. Of the 7 categorised thus, 5 stand out, 3 were labelled as part of the RAB GTPase superfamily. RAB GTPases are membrane localized cytosolic proteins that act as molecular switches, when bound to GTP they are able to recruit effector molecules which facilitate their many functions, most of those function are to do with the regulation of membrane trafficking and vesicle formation and function in multiple different organelles. One of these functions is to regulate, actin- and tubulin-dependent vesicle formation movement, and fusion with other membranes (Zhen and Stenmark, 2015). This may be notable as another category of proteins that were uniquely found with greater abundance in extract were 2 alpha tubulins and beta tubulins not found in GC. Perhaps pointing to the processes and movement of important vesicles that are key to *P. capsici* response to plant material. This could be to do with the release of effector molecule into the apoplastic space, or it could be to do with haustoria formation or another yet unknown mechanism of host adaption.

In addition to the RAP GTPases, two FKBP-type peptidyl-prolyl cis-trans isomerase molecular chaperones involved in protein folding were identified, this is in addition to the single Molecular chaperones HSP that was identified with increased intensity in extract. Like the HSP, the FKBP-type molecular chaperones supervise protein folding (Subin et al., 2016) and both types seem to be involved in the stress response, although each family is perhaps in response to a different kind of stress, be that temperature, salinity, ROS, or chemical. Or even just a natural proteins regulation pathway of host adaption.

4.4.5 Host to Host comparison

Examining host differences in abundancne of proteins in the early stages of infection can inform us of the host specific mechanism of host adaption and perception, and enlighten the study of broad host ranged pathogens like *P. capsici.*

A broad analysis of the patterns of host differences in protein abundance revealed that CE was able to give rise to a large induction of protein not found induced in *P. capsici* in TE. This similar pattern was also found in the RNA sequencing data in chapter 4. It was suggested there that this may be due to the fact that *P. capsici* is routinely cultured on V8 vegetable juice, the main component of which is tomato fruit juice. It could be speculated that this gives the Lab strain of *P. capsici* that is used for these experiments some long term or short term, metabolic or other adataption or preference for growing on tomato based media. The timeline of infection also progresses quicker on tomato plants perhaps revealing a greater adataption to tomato. What the proteomics does reveal which was not noted in the RNA sequencing results, however, is the relative similarity in proteins found differentially abundant in cucumber and GC

4.4.6 Protein function in Host to Host comparison

In tomato, only 11 proteins were found with increased intensity compared to cucumber. In terms of timing 5 were found only in one-time point 2 hpi and 4hpi, 3 were found at all time points and 3 additional proteins were found if you pooled all the time points. Suggesting close regulation of some proteins and overall need at all time points for others. There is no correlation however with function and at which time points these proteins are highly abundant. Notable for their consistently high abundance in TE compared to CE are three distinct regulator proteins. Annotated as 1) GTPase Rab2,

small G protein superfamily (Phyca11_504650) also induced in extract compared to GC, 2) Molecular chaperones HSP70/HSC70, HSP70 superfamily (Phyca11_104444) and 3) DEAH-box RNA helicase (Phyca11_538207) also induced in extract compared to GC. These three regulatory proteins may be key to distinct mechanism in tomato infection in *P. capsici* and thus warrant further examination.

Whereas the reaction of *P. capsici* to CE seems to mirror the proteins intensity pattern seen when comparing extract to GC, and in fact, the majority of the differences noted in extract to GC comparison are largely driven by the response to TE and not the response to CE. And the functionality and distribution of the quantities of those functional groups of those proteins with greater abundance in CE are also very similar to that of the proteins induced in GC. The two largest functional groups found with greater abundance in CE are metabolic proteins and then oxidoreductase/ mitochondrial proteins. Maintained in these metabolic proteins are members of amino acid synthase pathways, high energy carbohydrate biosynthesis pathways such as succinate, pyruvate, and malate, and fatty acid synthesis pathways. All pathways that were also found enriched in GC, many of these proteins that make up these pathways can be found at a higher intensity in GC than CE. CE also seems to induce a higher respiratory activity similar to GC and paired with the induction of metabolic activity. CE is also able to induce other similar proteins to GC, such as the two HSP that are found with increased intensity in CE and tRNA synthases. In fact the phenotype of GC and CE resemble the published information about the intensity patterns of germinating cyst in other *Phytophthora* species. TE on the other hand induced few proteins, but those that it does induce are interesting sources of further research.

4.4.7 Potential weaknesses of an *in vitro* model approach:

The obvious weaknesses of this model are the lack of natural plant tissue, and the physical effects of the plant cells and leaf structure on *P. capsici* during infection and invasion. There is also a timescale issue, in that cysts in nature would only be confronted with the surface of a plant, whereas the cysts in this experiment are immediately confronted with the plant cell debris and plant derived nutrients, perhaps not triggering the expected life cycles stages, i.e. it may trigger a more vegetative growth stage as appose to germination and infection. However, they are being challenged with the internal environment of the plant and many chemicals and metabolites that are unique to that host plant. The plant also has no possibility to mount an immune response, so the initial defence mechanism that may trigger some elements of the *P. capsici* lifecycle, that is, notable elements of ETS may not be triggered and represent a large amount of the chemical signals normally found upon plant invasion may be missing. It has also been noted that the metabolic system of *Phytophthora* species is also paired with that of the plant, in terms of the derivation of some required nutrients. The intimate chemical relationship between pathogen and host may not be completely recreated just with the plant extract. (Rodenburg et al., 2019).

One of the purposes of this study was to validate the RNA sequencing that was carried out in chapter 4. When comparing to the RNA sequencing data there are similarities and differences in what was observed. At a gross level the results from the RNA sequencing and the proteomics appear similar. In both data sets there was a lot more differentially expressed gene/proteins induced in GC sample than in extract. Likewise in the host to host comparison, there were a larger amount of differentially expressed gene/proteins induced by CE than TE. There are comparisons to be made in the

functionality of the genes identified, in both data set. For example, there were a large amount of oxidoreductase enzymes, and enzymes with metabolic function. One of the groups of genes that were found heavily up-regulated in the RNA sequencing in terms of function were transmembrane, or membrane localised protein, however in the proteomics whilst a few proteins that were found differentially regulated were localised to the membrane, there were not largely enriched, and many were predicted to be mitochondrial membrane proteins.

When examining the specific genes/proteins that were identified and differentially expressed in the proteomics and RNA sequencing however there is less alignment in the comparisons made. Of the differentially expressed gene in the RNA sequencing 407 of a total 1092 were found identified in the proteomics. And of those 407 the proteomics identified 55 of them to also be differentially expressed. When examining those 55 gene/proteins that both data set showed to be differentially regulated, few matched in which conditions they were found to differentially regulate in. This is likely down to two possible reasons, 1) that there is poor repeatability within extract inoculation experiment, and although pains were taken to separate biological replicates, plants used in say the proteomic experiment were likely of the same batch grown. These plants therefore are more of a similar age and grown at a similar time of year and more similar condition, those plants from the batch used for the RNA sequencing experiment which would be more similar to each other that the proteomics. Additionally, although separate subcultures of *P. capsici* were used for each biological replicate, it is unclear how much *P. capsici* alters through sub-culturing and on a yearly cycle, the subculture for each experiment would have only been separated a week or so prior to the experiment meaning they again may be more similar than the same subculture of *P. capsici* a year later. And 2) there may be a large difference in the

mRNA transcript found in *P. capsici* and the protein content, this is, the perineal problem of using transcriptomics to assess protein content and cellular phenotype thereof.

4.4.8 Conclusion and further work

The purpose of this proteomics based study was, initially, to further validate and explore the extract inoculation assay as a model for “omics” analysis of the yet understudied pathogen response to plant host during the infectious lifecycle stages. As such, the data shown in this chapter has raised some interesting questions and avenues for further research, however it is yet unclear if the extract inoculation assay is a repeatable and accurate model of infected plant leaves.

This study along with the RNA sequencing in Chapter 3 have pointed to the induction of molecular processes that may be important for germinating cysts, they have also shown that host extract is able to induce changes in transcript and protein levels when compared to those found in germinating cysts incubated without extracts. The broad context patterns of that gene/protein expression that can be found in both studies suggest key molecular processes in host adaption. It has been clearly demonstrated that large changes in the protein suite with putative roles in the metabolism of *P. capsici* occur during cyst germination and when confronted with host extracts. Also notable are the changes in proteins with potential roles in protein synthesis, modification and metabolism, high energy carbohydrate biosynthesis, and fatty acid biosynthesis. It is also now clear that a less dramatic change in protein complement is necessary for *P. capsici* to adapt to tomato extract, this could be due to several reasons, perhaps to that fact that *P. capsici* is routinely cultured on V8 juice, the prime ingredient of which is tomato fruit juice, in the laboratory, or perhaps because tomato

is more susceptible to *P. capsici* infection. It is clear that the same mechanisms that are being utilized by *P. capsici* to successfully colonise tomato, as suggested by the protein/gene expression patterns shown here and in chapter 3, are not the same as those used to colonise cucumber, as demonstrated by extract inoculation at least.

One interesting finding of these studies is the importance of the mitochondria and the potential changes in the respiratory capacity to cyst germination, and the adaption to certain host plants i.e. adaption to CE. This study has also identified a potential group of genes that are important for adaption to host plants, in particular to TE. The RAP GTPase superfamily warrants further study for its effect in host adaptation in *P. capsici*. The Heat shock protein chaperone family also seems to have a notable role in cyst germination even response to host extracts.

This study and the study in chapter 3 has potentially identified some general mechanisms and protein families that may have a role in host adaptions, and thus potential avenues for further study into the dynamic host adaption mechanisms of *P. capsici*

Chapter 5: Attempts to Implement the CRISPR-Cas9 system for gene knock out in *P. capsici*

This work includes contributions from Rudd Grootens and Joram Westera

5.1.1 Introduction

Phytophthora capsici is a broad range oomycete plant pathogen able to cause disease on a wide range of economically important crop plants. It is well placed to become a model organism for the study of oomycete genetics and molecular host range determinants. However in *P. capsici* as well as other *Phytophthora* species, efforts to functionally analyse genes of interest for their effect on host determination, pathogenicity, or life cycle, has been hampered by these species' limited genetic tractability. Multiple transformation procedures have been published for many *Phytophthora* species. Four methodologies have been utilized for the transformation of *Phytophthora* species, these include PEG/calcium mediated protoplast transformation, zoospore electroporation, micro projectile bombardment, and agrobacterium mediated transformation (Fang and Tyler, 2016, Birch and Whisson, 2001, Judelson, 1997), however these methods vary in efficacy and reproducibility, and the success of these different approaches varies from species to species. Efficient transformation protocols are necessary to establish the genetic tractability of an organism. Although gene insertions, over expressions and knock-ins have long been possible, this only gives us a small picture of the role of a gene of interest. Depending on the transformation technique, the nature of random genome integration and the genetic plasticity of the species being studied, the efficiency, stability and even the phenotype of transformants can vary (Fang et al., 2017). In situ silencing would, in contrast, represent an additional and alternative approach to study gene functionality.

RNA interference (RNAi) has been utilized (Ah-Fong et al., 2008, Wang et al., 2011, Whisson et al., 2005) for functional studies. However knock-downs are not complete and the effectiveness varies between genes, experiments and laboratories. Although the more studied *Phytophthora* species have a growing number of methodologies for the gene insertion and deletion, the effectiveness of these methodologies and the reproducibility varies.

The last few years have seen a revolution in the molecular biology field, with the advent of the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 genome editing technique. The first CRIPSRs were identified in the late 1980s and were soon understood to be part of a bacterial anti-viral defence system, with the later discovery of the associated Cas proteins. It was not until 2002 (Hsu et al., 2014) that two groups published genome editing techniques in mammalian cell using the CRISPR/Cas9 system. The system utilizes the Cas9 protein which is bound to a 20 base pair long protospacer sequence, the gene specific portion of the single guide RNA (sgRNA) which enables directed and specific knock-outs or edits to genes of interest. The addition of a DNA repair template is optional and allows the introduction of specific mutations. Once guided to the site of mutation the Cas9 makes a double stranded cut in the DNA: the cut occurs at the protospacer adjacent motif (PAM) sequence, a 3 base pair long sequence (5'-NGG-3') which is specifically recognized by the Cas9 protein and directs the nuclease double stranded cut. The endogenous repair mechanism can introduce mistakes when joining the two DNA ends back together, these mistakes are referred to as non-homologous end joining (NHEJ) and are the mechanism by which mutations, deletion or insertion that can cause frameshifts resulting in gene knock outs. Recently a methodology for the genome editing of *Phytophthora sojea* using the CRISPR/Cas9 system was developed (Fang

and Tyler, 2016). This was an important breakthrough in the molecular investigation of oomycete plant pathogens. However, it seems that every oomycete species represent a new challenge in implementing this tool. Following the publishing of genomic editing using CRISPR/Cas9 in *P. sojea,* several *P. sojea* genes have been mutated, edited or knocked out (Ma et al., 2017, Fang et al., 2017). More recently a procedure for the functional editing of key genes has also been published in *P. capsici* (Wang et al., 2018) and it has been used to investigate the effect of specifically inserted single point mutations in the Oxathiapiprolin resistance gene ORP1 (Miao et al., 2018). Indeed, a mutated version of this gene (PcMuORP1) has since been shown to have utility as a selection marker for use with *P. capsici* and the CRISPR/Cas9 system (Wang et al., 2019). Functional studies of two genes from *Phytophthora palmivora* that have roles in pathogenicity have also been carried out (Gumtow et al., 2018, Pettongkhao et al., 2020). Additionally, several systematic attempts to implement the CRIPSR/Cas9 system in *P. infestans* (van den Hoogen and Govers, 2018) have had no success. It seems the success of the Cas9 methodology may be very species or even strain dependant. The dynamic genome of most Phytophthora species may also limit the length of time for which these mutation have influence. But for some species the Cas9 system has been able to increase genetic tractability and allow better investigation of the functional roles of genes of interest.

This project aimed to develop and utilize the CRISPR/Cas9 gene editing and knockout system in *P. capsici,* building on previous studies in this area. Here the aim was to use a single sgRNA system to enable quick, efficient total knock out of key genes. Previous studies have identified multiple genes of interest that may have functions in the host range determination in *P. capsici*. To truly investigate the determinants of broad host range it would be useful to be able to analyse the host pathogenicity phenotype of *P.*

capsici lacking these genes of interest. Many genes that have been identified because of their differential expression in the infection of different host plants are effector genes, more specifically they are RXLRs, a conserved family of effectors. In fact RXLRs are some of the most differentially expressed genes in these experiments (Chapter 1 Unpublished Huitema Lab). However, it is unknown if these effectors have host specific phenotypes. A methodology for quick and efficacious knock out of these sorts of genes would be a boon for the molecular investigation of *P. capsici*.

The methodologies that are described in this study are adapted from those published by Fang and Tyler (2016), and Fang et al. (2017). Transformation protocols have been adapted from those published by Wu et al. (2016) and Wang et al. (2018). Initially two transformation protocol were utilized: agrobacterium mediated transformations (AMT) and PEG/calcium. AMT should be a more robust transformation with potentially multiple copies of the transgenes being incorporated into the genome. However, due to low transformant yield, PEG/calcium transformations were explored as they resulted in a greater number of transformants. Although these methodologies underwent several rounds of optimization, the Cas9 mediated gene editing of *P. capsici* genes published by others (Miao et al., 2018, Wang et al., 2019) was not able to be replicated. The experimental efforts to implement the CRIPSR/Cas9 system for gene knock down are described here, along with suggestions for future work in order to successfully implement this technology in the broad range plant pathogen *P. capsici*.

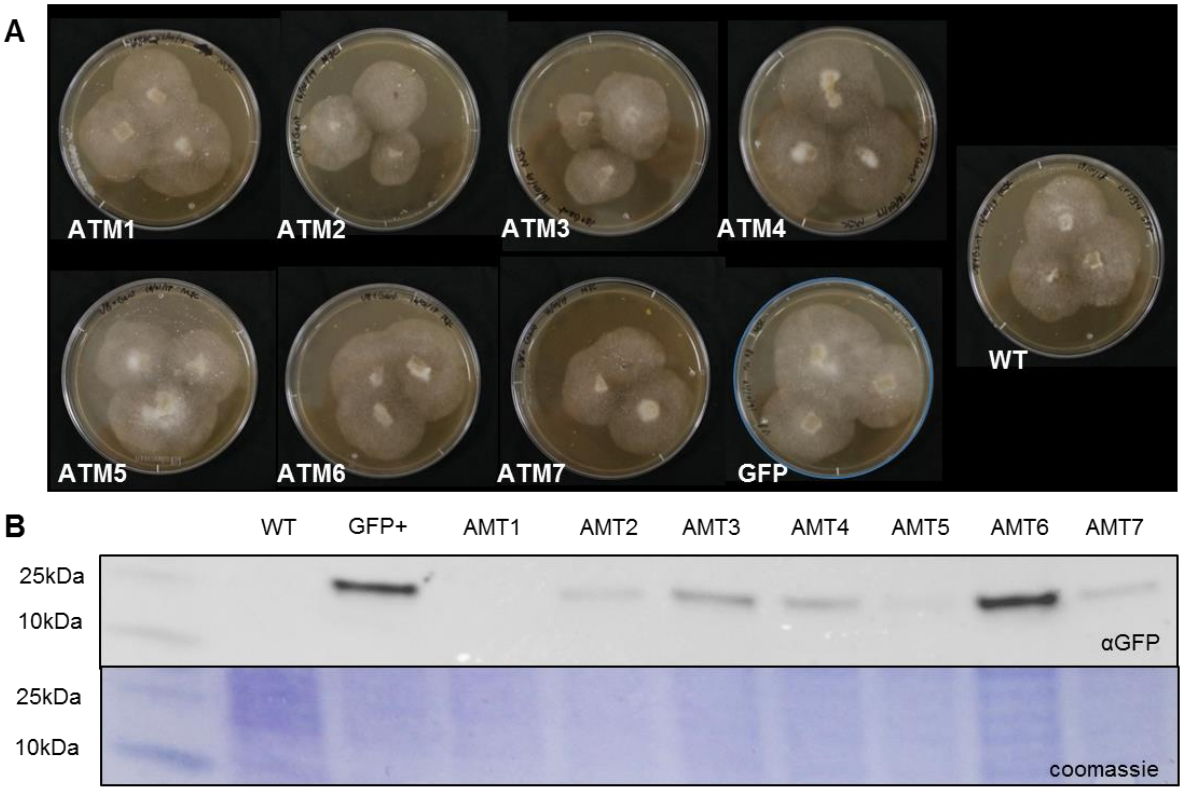


Figure 5.1: 7 AMT Transformants.

GFP expression was assessed in all seven clones (Agro-mediated-transformation [AMT]1-7) and compared to a GFP+ clone transformed by other means, and a wild-type (WT) control. Transformants and controls grown on V8 agar (A). Western blot shows the 7 transformants and controls total protein extract immunoblotted with anti-GFP and visualised with a secondary antibody conjugated to HRP. The immunoblot is shown stained with coomassie to show correct loading and total protein

5.2 Results

5.2.1 AMT transformation

The primary vector for AMT was pCB301TOR, originally developed for transformation of the oomycete plant pathogen *Phytophthora palmivora* (Wu et al., 2016). It was constructed through the combination of the pC301 backbone developed for agro-infiltration and plant transformation, including the T-DNA left and right boarders (LB and RB), and the oomycete expression vector pTOR. pCB301TOR itself contains a HAM34 promotor-terminator flanked MCS and a hsp70 promoter-NPTII-hsp70

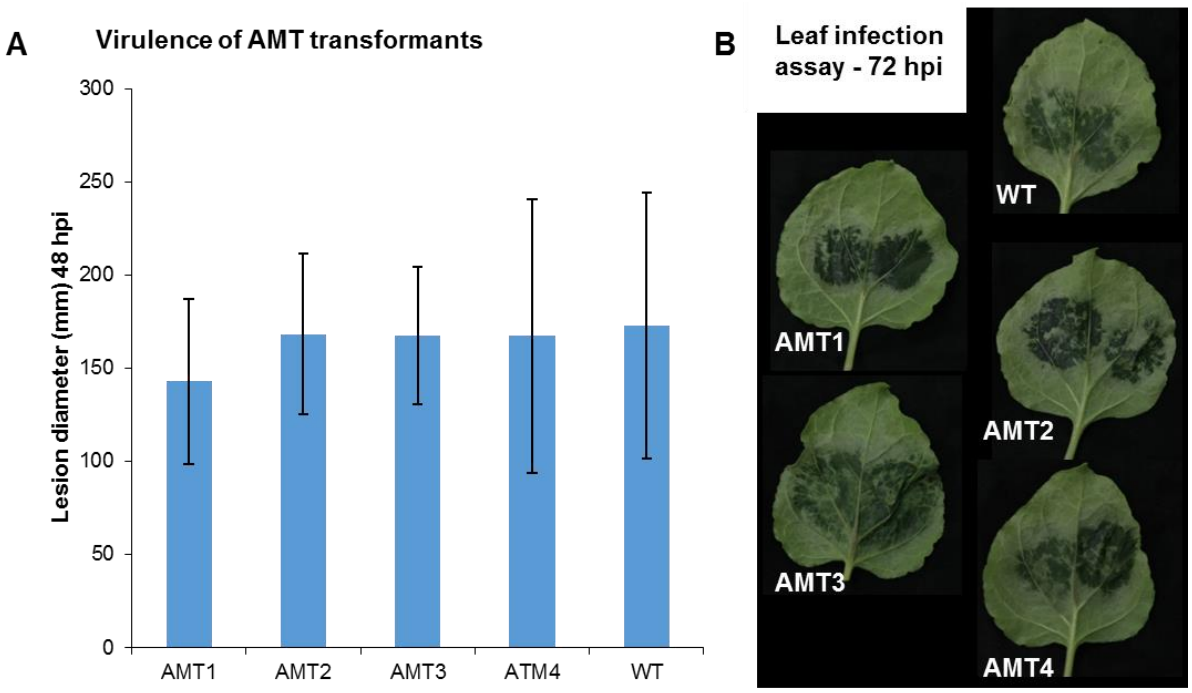


Figure 5.2: Transformants AMT1-4 leaf infection assay.

Four transformants from the AMT and a *P. capsici* wild type (LT1534) strain, were then assessed for virulence in *N. benthamiana* leaves. 10 µL of zoospore suspension was placed on the abaxial side of a *N. benthamiana* leaf, lesion diameter in mm was measured at 48 hpi (error bars show standard deviation) (B) and pictures of infection taken at 72 hpi 2 large water soaked lesion on the left and right hand side of the leaf can be seen surrounding the initial inoculation point (C). Two tailed unequal variance student t-test was conducted, significant bonferroni adjusted P values are shown. Adjusted P values for all comaprisons between all samples were >0.05 and therefore non-significant, not shown on the graph.

terminator for selecting transformants using the antibiotic G418. For testing of this vector in *P. capsici,* eGFP was cloned into the MCS of pCB301TOR. AMT transformations were conducted with pCB301TOR-eGFP, to examine the viability of using AMT in *P. capsici*. 7 transformants were produced by a single transformation experiment (Figure 5.1 A). These were then tested for GFP expression by western blot. Expression levels varied, however it appeared that all but one of the transformants showed expression of GFP (Figure 5.1 B). To check the viability of AMT as a new transformation method it was also necessary to check that transformants maintained pathogenicity and virulence. Leaf infection assay was carried out in *N.*

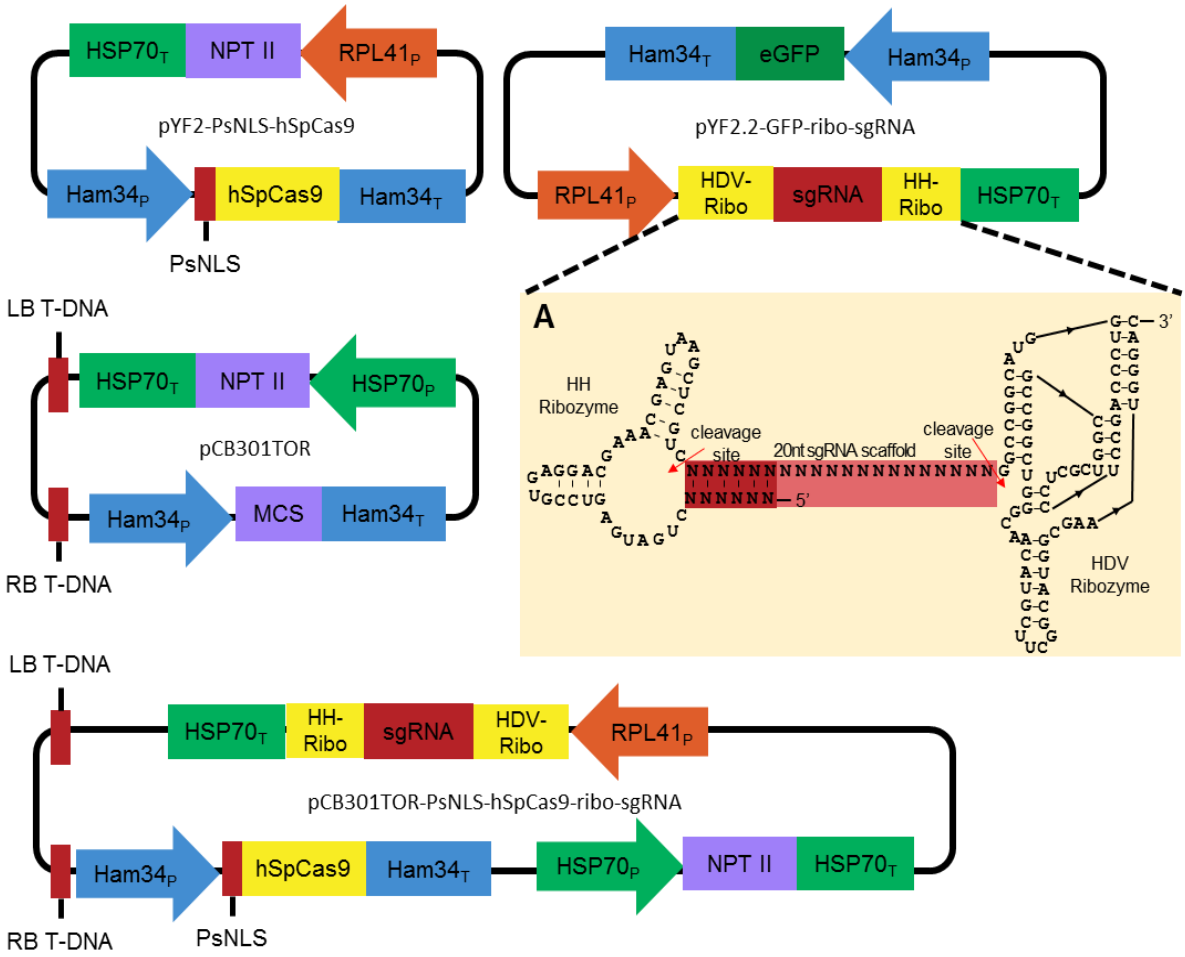


Figure 5.3: the template for the bespoke dual Cas9 and sgRNA vector, and the 3 vectors used in its construction.

pCB301TOR is the backbone, used of AMT, containing the promotors and terminators for *Phytophthora* (HSP70 and Ham34) driving expression of a multiple cloning site (MCS) and Neomycin Phosphotransferase II (NPT II) resistance gene. The other two plasmid used to make pCB301TOR-PsNLS-hSpCas9-ribo-sgRNA were pYF2-PsNLS-hSpCas9 containing the human codon optimised *Streptococcus pyogenes* Cas9 protein (hSpCas9) *P. sojea* nuclear localisation signal (PsNLS) fusion protein, and pYF2.2-GFP-ribo-sgRNA containing the sgRNA cassette. The sgRNA cassette in more detail contains two ribozyme elements up and down stream of the 20 nucleotide (nt) sgRNA, the HH ribozyme and the HDV ribozyme, both of which are cleaved off during transcription (FANG 2016). For the construction of the dual vector the PsNLS-Cas9 was placed in the MCS of pCB301TOR and the sgRNA caste with RPL41 promotor and HSP70 terminator was placed between the selection gene and the left boarder (LB) T-DNA .

benthamiana. Lesion size was measured at 48 hours post inoculation (hpi). There is no notable difference in between wild type (WT) *P. capsici*, LT1534 and the four transformants tested (Figure 5.2).

#	Name	Sequence	Target	Cut Site (BP from start)
1	ABC1-Δ668	GATATCGTCATACAGCAGCTCGG	Phyca11_6218 (ABC1)	668
2	RXLR001-Δ160	CCACCCGTTGACTTTCTTCCAGG	Phyca11_21535 (PcRxLR001)	160
3	RXLR135-Δ99	TCACTGGAAAGGACTTGGTTTGG	Phyca11_534478 (PcRxLR135)	99
4	RXLR374-Δ13	AGCAATTATGAGGAGAGTGTAGG	Phyca11_13953 (PcRxLR374)	13

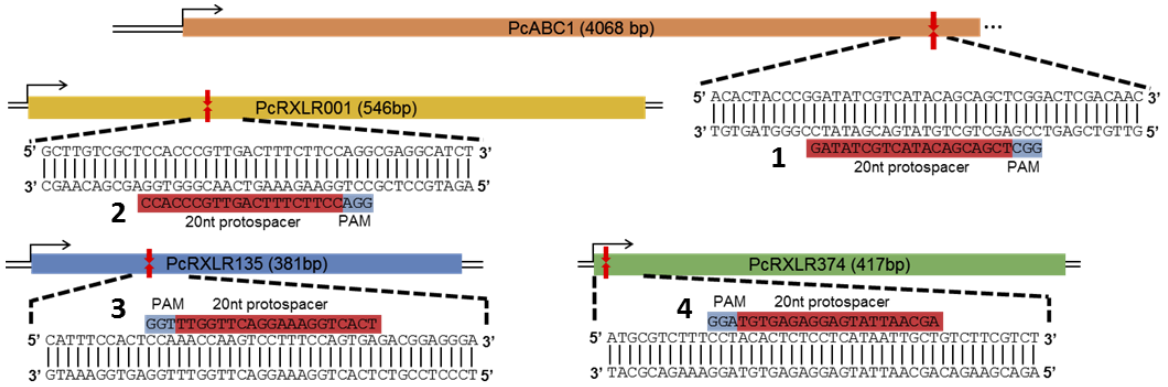


Figure 5.4: sequence and schematics of 4 of the sgRNAs designed and transformed.

The sequence, target gene and cut site of each of the sgRNAs that were transformed in *P. capsici* is shown in the table. The target genes and the respective cut site in relation to the rest of the gene, the strand they target (5' or '3) and the PAM site for each is shown in the schematic underneath.

5.2.2 Construction of dual Cas9 – sgRNA vector for AMT

Due to the low transformation efficiency of AMT, the transformation of two plasmids was viewed to be unfeasible. In order to insert both the Cas9 transgene and a sgRNA cassette into *P. capsici* for use in gene knock-out experiments, it was deemed necessary to produce a single dual plasmid with both the Cas9 gene and the sgRNA cassette. The pCB301TOR plasmid was used as the backbone. The two plasmids pYF2.2-GFP-ribo-sgRNA and pYF2-PsNLS-hSpCas9 were used as the source of the Cas9 and the sgRNA (Fang and Tyler, 2016) (Figure 5.3). This enabled amplification and cloning of a *Streptococcus pyogenes* Cas9 gene, encoded by codons optimized for human. This Cas9 version has been shown to function in diverse organisms (Peng et al., 2014, Zhang et al., 2014, Liu et al., 2011) with a 5' *P. sojea* nuclear localisation signal (NLS) necessary for more efficient localisation of the Cas9 protein into the nucleus (Fang and Tyler, 2016). The sgRNA cassette was also altered for

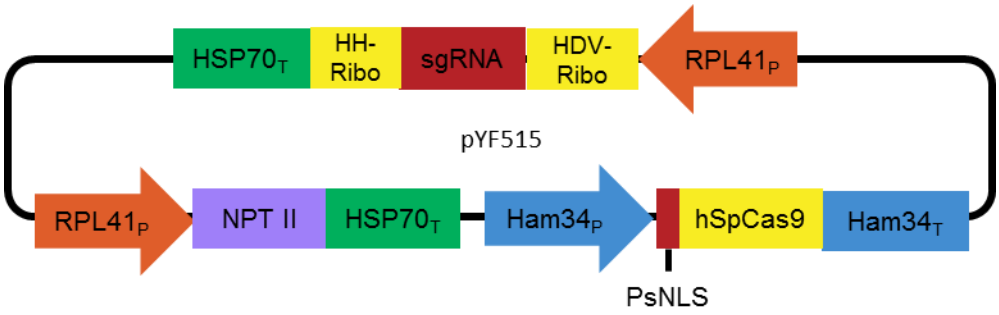


Figure 5.5: All in one plasmid pYF515 from Fang et al 2017.

A shematic showing the dual Cas9, sgRNA cassette pYF515. it also contains the Neomycin Phosphotransferase II (NPT II) resistance gene, all genes under their respective promotors and terminators.

Phytophthora biology, in most organisms the sgRNA is synthesised by RNA Polymerase III, however there is no RNA polymerase III found functional in *Phytophthora* species to date, so a different method for sgRNA synthesis was necessary. Flanking the sgRNA with two cis-acting ribozyme sequence downstream of a *P. sojea* RPL41 promotor has been shown to generate the sgRNA in a *Phytophthora* system (Fang and Tyler, 2016) (Figure 5.3). These two cassettes, NLS-hSpCas9 and HDV-ribozyme-sgRNA-HH-ribozyme, from these two plasmids were combined into one, using the pCB301TOR AMT backbone named pCB301TOR-PsNLS-hSpCas9-ribo-sgRNA (Figure 5.3).

5.2.3 sgRNA design and AMT Cloning

Initially a single sgRNA was cloned for the ABC transporter PcABC1 (Phyca11_6218) here named ABC1-Δ668 (Rudd Grottens) (Figure 5.4). This was successfully cloned into pCB301TOR-PsNLS-hSpCas9-ribo-sgRNA. AMT transformation was conducted and 6 transformants created. Successful transformation and expression of transgenes was analysed using PCR of genomic DNA and RT-PCR of total RNA (Figure 5.5). The expected size of the NPTII amplicon was 771 bps, although an unspecific band appears in the WT and transformant 1, all transformants display NPTII integrated into

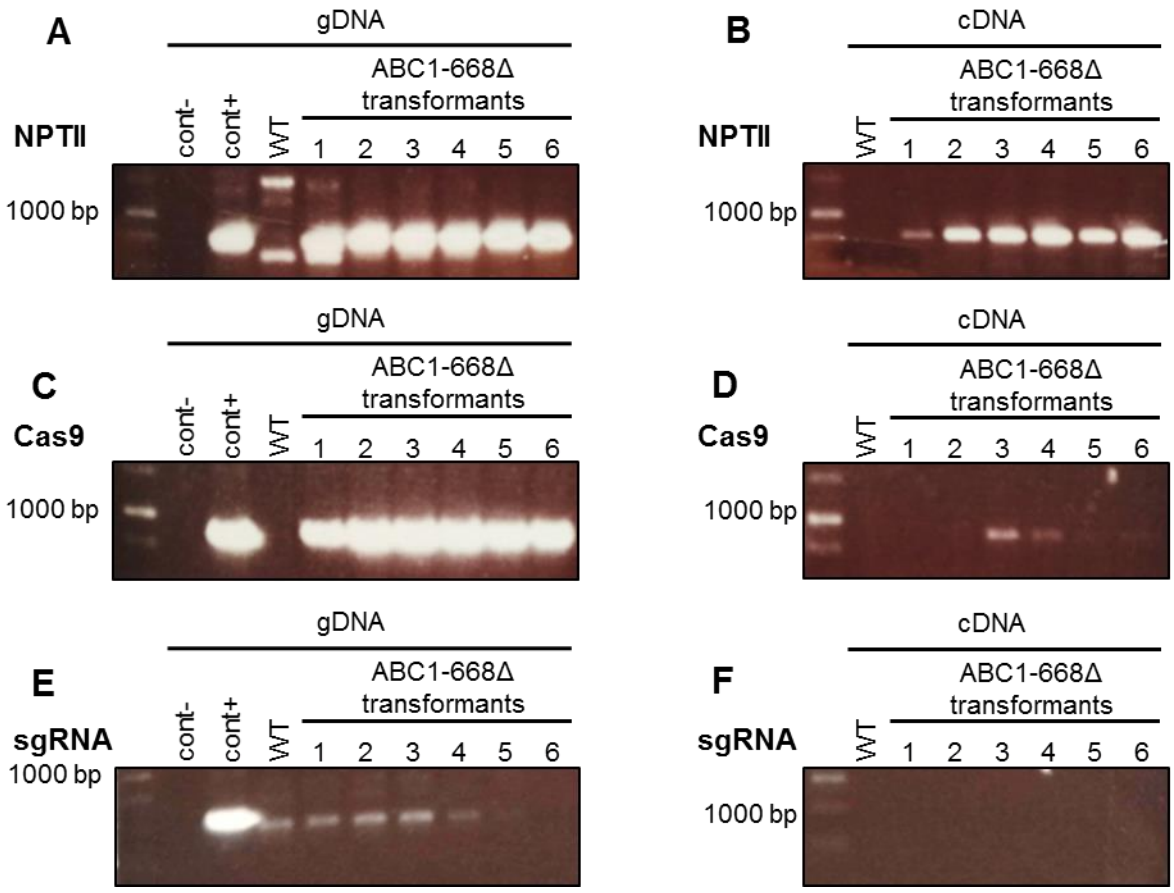


Figure 5.6: PCRs of AMT transformants with Cas9 and sgRNA ABC1-668Δ.

Genomic DNA and total RNA (from which cDNA was derived) was extracted from Wild-type *P. capsici* (LT1534) and the 6 transformants derived from AMT (ABC1-668Δ transformants 1-6). To test the success of the transformation and PCR of the selection gene NPTII was conducted from gDNA (A) and cDNA (B) with an expected amiplicon of 771 bp. PCR amplification of both the Cas9 protein in both gDNA (C) and cDNA (D), and of the sgRNA cassette in both gDNA (E) and cDNA (F), the expeted size of these amplicon was 883 bp and 554 bp respectively. In all cases 1000bp maker of the DNA ladder has been indicated. The positive control for all was a DNA template of the dual plasmid pCB301TOR-PsNLS-hSpCas9-ribo-sgRNA, and a negative control of water.

gDNA and RNA transcript expression (Figure 5.5 A-B). Of course NPTII expression is necessary for growth on G418 media, on which the transformants were cultured and selected, of more relevance is the expression of Cas9 and the sgRNA. PCR of the Cas9 gene, with an expected amplicon of 1188 bps showed good expression, albeit with a band slightly lower than expected (Figure 5.5 C). Additionally, only two transformants showed any substantial expression in transcript form (Figure 5.5 D).

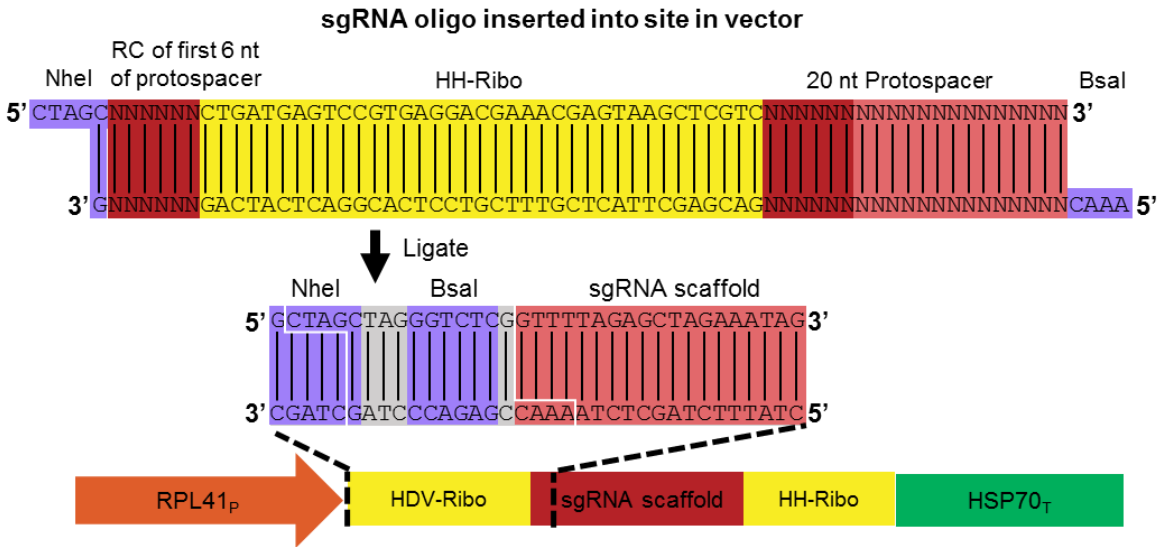


Figure 5.7: Schematic for the insertion of the sgRNA into the sgRNA cassette with ribozymes.

For sgRNA design two annealing oligos containing HH-ribozyme and the 20 nucleotide protospacer, which an addition 6 nt of the protospacer which are used to create the secondary structure of the ribozyme. This can then be ligated into the vector using the NheI and BsaI restriction sites. Protocol is as described by Wang et al. 2018.

The sgRNA, however, displayed no integration into the gDNA and no expression in RNA form (Figure 5.5 E-F).

5.2.4 Protoplast transformation and CRISPR element expression

Due to low transformant numbers, lack of expression of key transgenes from the vector and CRIPSR/Cas9 system, and the publication of new methodologies, a new methodology for *P. capsici* transformation was trialled. This new transformation method was made more feasible by the development of a new dual Cas9, sgRNA plasmid pYF515 for use in *Phytophthora* (Fang et al., 2017) (Figure 5.6) An altered version of the protoplast transformation described by Wang et al. was developed (Wang et al., 2018); The recovery stage was maintained, however the growth stage was altered to only include liquid medium growth stages, and more than sufficient numbers of protoplast were obtained from this growth protocol.

Table 5.1: Summary of transformant testing.

To test if the transformants were positive PCR of the Cas9 transgene was performed form gDNA. The number of positive and negative transformants are shown for each sgRNA vector that was transformed. To test the success of the CAS9-sgRNA complex in mutating the target gene, the target gene was sequenced, the number of transformants that were sequenced and at how many weeks post transformation (wpt) the sequencing was carried out is shown in the lower portion of the table, in addition to the number of transformants found with or without mutations.

	Transformants		
PCR for Cas9 Expression	RXLR001-160Δ	RXLR374-13Δ	RXLR135-99Δ
+	6	5	9
-	7	21	22
	Transformants		
Target Gene Sequencing	RXLR001-160Δ	RXLR374-13Δ	RXLR135-99Δ
4 wpt	0	5	7
5 wpt	0	21	23
10 wpt	0	5	7
With Mutation	0	0	0
Without Mutation	0	31	37

For this new plasmid and transformation system, 3 new sgRNAs were designed. These new sgRNA were designed for 2 RXLRs that were shown (Chapter 1) to be uniquely up-regulated in cucumber leaf infection, out of 4 host plants tested (PcRXLR135, and PcRXLR374). The third new sgRNA was designed to target another RXLR, PcRXLR001. These were named RXLR001-Δ160, RXLR135-Δ99, RXLR374-Δ13 (Figure 5.4) and inserted into pYF515 (Figure 5.7) and then transformed into *P. capsici* using the protoplast transformation method. From the three plasmids transformed, here referred to by which sgRNA they contained, RXLR001-Δ160, RXLR135-Δ99, RXLR374-Δ13, 13, 26 and 31 transformants were obtained for each plasmid respectively (Table 5.1). They were collected in two “batches”: the first were initial colonies emerging 2-3 days post transformation, first to grow through the V8 agar selection disc placed atop the PAM recovery media the day following

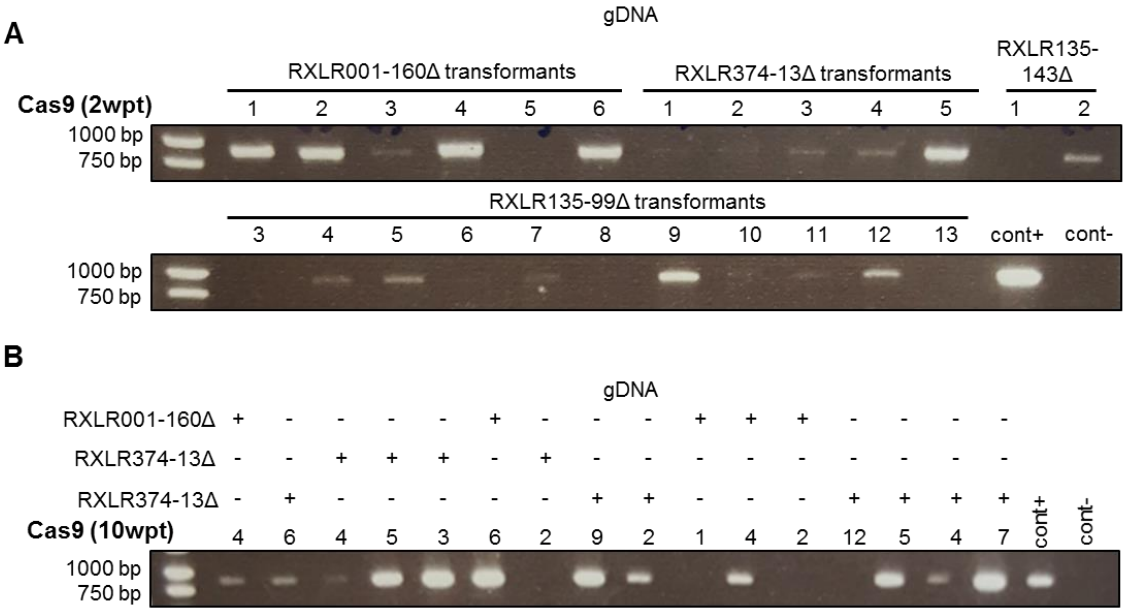


Figure 5.8: Cas9 expression PCRs in Protoplast transformants gDNA.

Genomic DNA from several transformants derived from protoplast transformation, 6 from the transformation of the dual Cas9 sgRNA plasmid pYF515 containing the sgRNA RXLR001-160Δ, 5 from the sgRNA RXLR374-13Δ, and 13 from the sgRNA RXLR374-13Δ. PCR amplification of the Cas9 transgene was conducted 2 weeks post transformation (wpt) and again at 10 wpt. Only 16 of the transformants were tested at 10 wpt, which sgRNA was transformed into them, and which number transformant they are displayed above the 10 wpt gel picture. The expected Cas9 amiplicon was 771 bp, and on all gel images the 1000 bp and 750 bp marker of the DNA ladder have been indicated. In all cases the positive control was a DNA template of the dual plasmid pYF515 and the negative control was water.

transformation. The second batch were the many more colonies that emerged on days 4-6 post transformation. Of the initial colonies (of which there was 6, 5, and 13 from RXLR001-Δ160, RXLR135-Δ99, RXLR374-Δ13 respectively), when gDNA was tested for the presence of the transgene NLS-Cas9 all 6 RXLR001-Δ160 transformants, all 5 RXLR135-Δ99 transformants, and 9 of the 13 RXLR374-Δ transformants were positive (Figure 5.8 A). 3 RXLR135-Δ99 positive transformants and 2 each of the RXLR001-Δ160, RXLR374-Δ13 positive transformants were examined using RT-PCR, cDNA derived from total RNA was used to probe for the presence of sgRNA and Cas9 transcripts. All of the tested samples were positive for sgRNA and Cas9 transcripts (Figure 5.9).

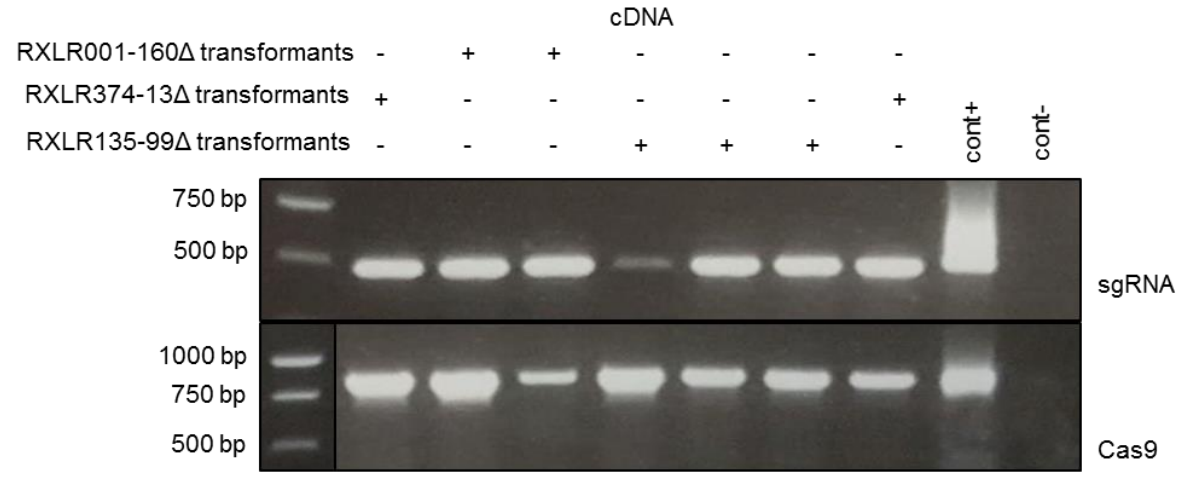


Figure 5.9: Cas9 expression PCRs in Protoplast transformants cDNA.

Total RNA from several transformants derived from protoplast transformation, 2 from the transformation of the dual Cas9 sgRNA plasmid pYF515 containing the sgRNA RXLR001-160Δ, 2 from the sgRNA RXLR374-13Δ, and 3 from the sgRNA RXLR374-13Δ. PCR amplification from cDNA derived from these samples of the transgene Cas9 and the sgRNA cassette. For each sample which sgRNA was transformed into them, and which number transformant they are displayed above the gel picture. The expected Cas9 amiplicon was 771 bp and the expected sgRNA cassette amplicon was 554 bp, and on all gel images the 1000 bp, 750 bp and 500 bp marker of the DNA ladder have been indicated. In both cases the positive control was a DNA template of the dual plasmid pYF515 and the negative control was water.

From these samples, i.e. 3 RXLR135-Δ99 positive transformants and 2 each of the RXLR001-Δ160, and RXLR374-Δ13 positive transformants, total protein was also extracted and run on a SDS-page gel. This was blotted for the presence of the NLS-Cas9 fusion protein. It appeared that none of the positive transformants carried a full length Cas9 protein. However, samples loaded on the gel did seem to contain multiple smaller non-specific bands that were not present in the WT sample (Figure 5.10).

In addition to these tests, gDNA was probed again by PCR 10 weeks post transformation in the initial “batch” of positive transformants. The PCR for the transgene NLS-Cas9 showed that the majority of tested mutants maintained Cas9 expression. However, it appeared that 2 RXLR001-Δ160 transformants (numbered 1 and 2) and a single RXLR135-Δ99 (numbered 12) had lost expression (Figure 5.8).

5.2.5 Sequencing possible Cas9 mutants

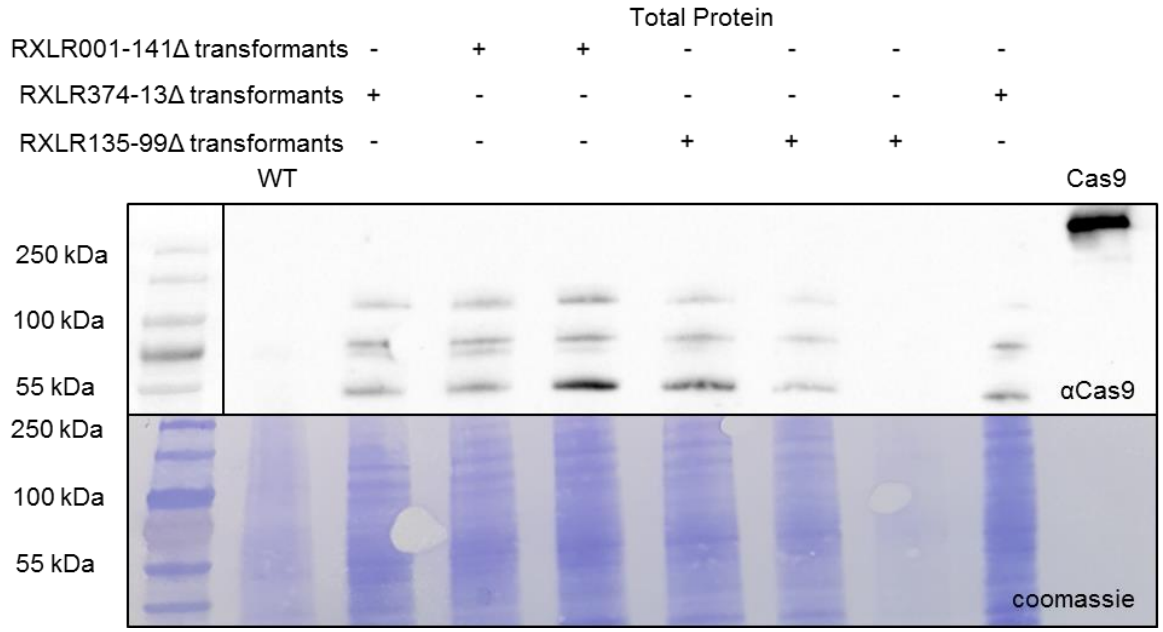


Figure 5.10: Cas9 expression in total protein derived from Protoplast transformants.

Total protein from several transformants derived from protoplast transformation, two from the transformation of the dual Cas9 sgRNA plasmid pYF515 containing the sgRNA RXLR001-160Δ, two from the sgRNA RXLR374-13Δ, and three from the sgRNA RXLR374-13Δ and wildtype *P. capsici* (LT1534) was collected. Western blot analysis for the Cas9 protein was conducted. For each sample, the respective sgRNA and transformant number is displayed above the blot. Total protein was blotted with anti-Cas9 antibody and visualised with a secondary antibody conjugated to HRP. The membrane was then additionally stained with coomassie. The positive control was Cas9 Nuclease, *S. pyogenes* (NEB).

Both the initial “batch” of positive transformants and the later “batch” of negative transformants were sequenced to see if any mutation in the target genes had occurred. Unfortunately it was not possible to amplify the target gene RXLR001 from any of the transformants or even the WT LT1534 strain of *P. capsici.* Therefore, only RXLR135-Δ99, RXLR374-Δ13 target genes were amplified and sequenced. Additionally the target genes of the positive mutants were sequenced twice, once at four weeks post transformation and once at 10 weeks post transformation. Four week post

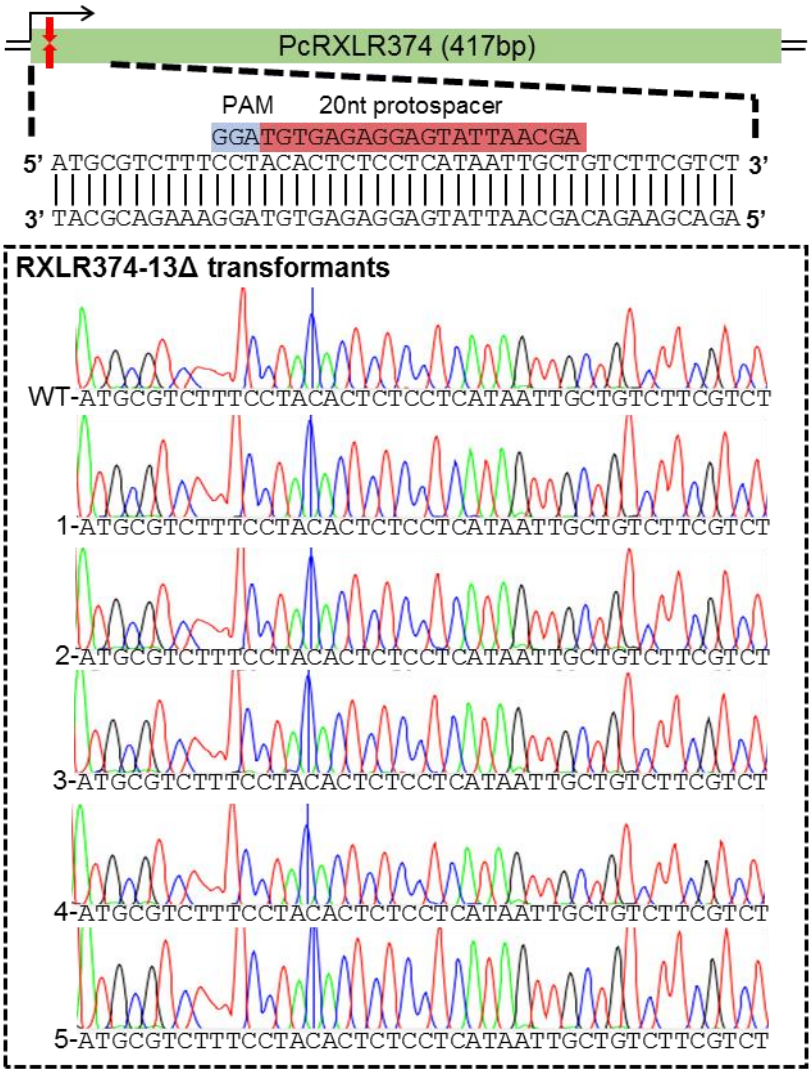


Figure 5.11: Sequence traces of PcRXLR374 form strains transformed with sgRNA RXLR374-13Δ.

From 5 transformants and a wildtype strain of *P. capsici* (LT1534) the PcRXLR374 gene was PCR amplified and sequenced. A schematic of the gene and the area where the sgRNA binds and cuts is show above. The 20 nt protospacer and PAM site are shown in red and blue respectively. The cut site is indicated on the gene by two red arrows. Sequence traces from the area around the cut site from the WT strain and the 5 transformants are shown below.

transformation of the five RXLR374-Δ13 transformants sequenced all showed no evidence of mutations around the PAM site indicative of Cas9 activity (**Figure 5.11**). Likewise at four weeks post transformation the 7 RXLR135-Δ99 transformants also showed no mutations (Figure 5.12). At 10 weeks post transformation sequencing the same transformants gave the same results, no mutants were derived from these transformation attempts.

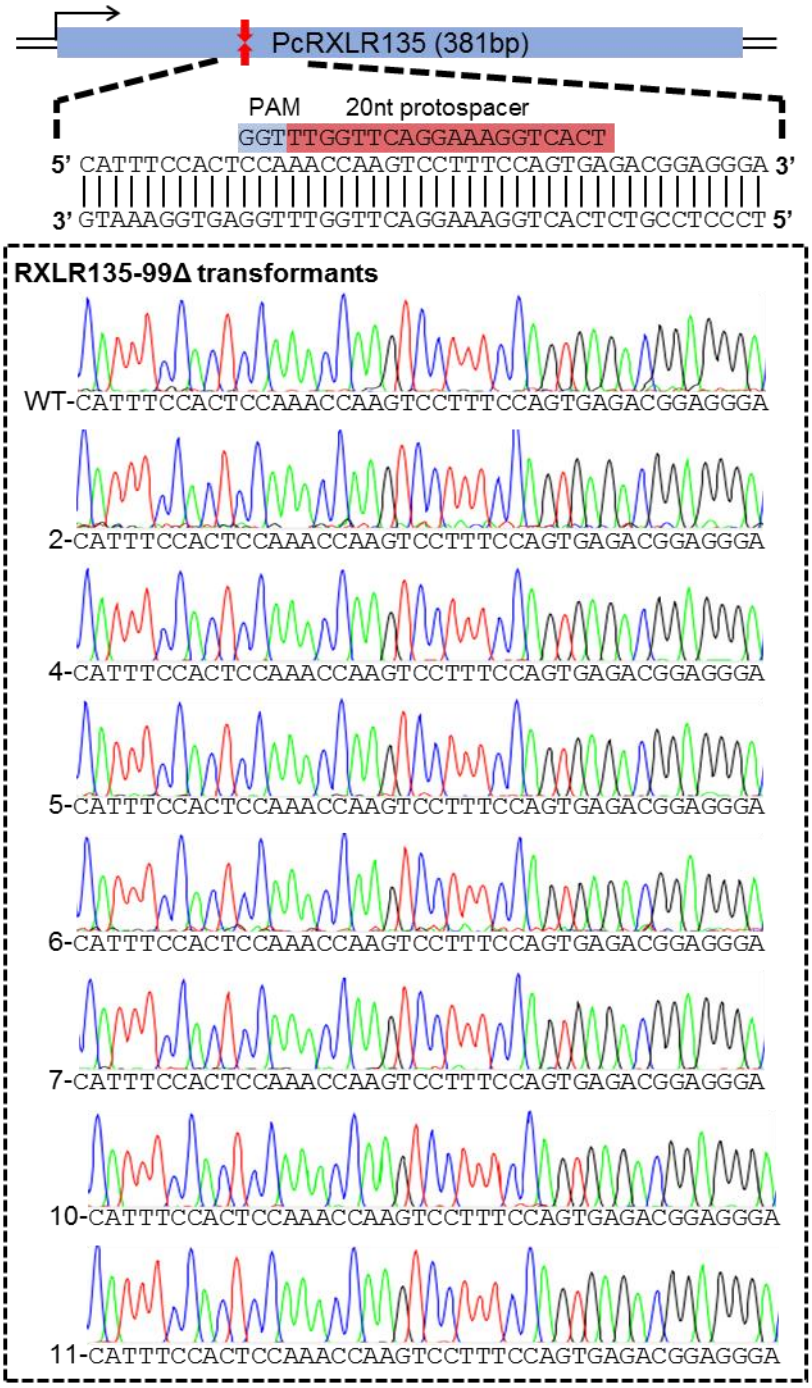


Figure 5.12: Sequence traces of PcRXLR135 form strains transformed with sgRNA RXLR374-13Δ.

From 7 transformants and a wildtype strain of *P. capsici* (LT1534) the PcRXLR135 gene was PCR amplified and sequenced. A schematic of the gene and the area where the sgRNA binds and cuts is show above. The 20 nt protospacer and PAM site are shown in red and blue respectively. The cut site is indicated on the gene by two red arrows. Sequence traces from the area around the cut site from the WT strain and the 7 transformants are shown below.

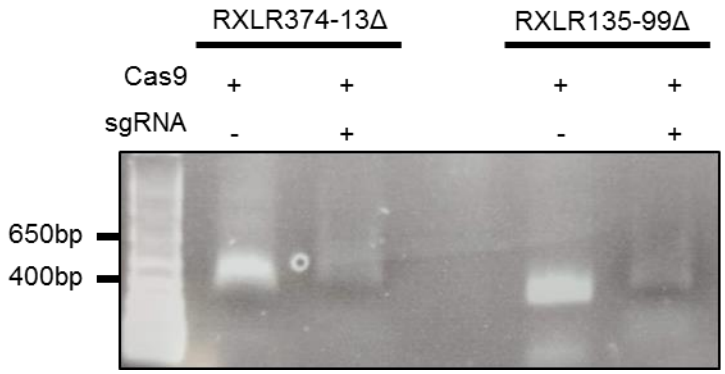


Figure 5.13: *In vitro* cleavage assay of the two transformed RXLR374-13Δ RXLR135-99Δ.

A gel showing the digested products of the PCR products of PcRXLR374 and PcRXLR135 following *in vitro* incubation with the Cas9 nuclease and their respective sgRNA. Undigested controls without sgRNA presence are also shown. 650 bp and 400 bp DNA band in the ladder are indicated. Expected full length of the PCR product of RXLR 374 is 417bp, and of RXLR135 is 386bp. Cuts sites as directed by the sgRNA for RXLR374 is 13bp from the start of the sequence, and for RXLR135 it is 99bp from the start of the sequence.

5.2.6 Further sgRNA development; *in vitro* transcription cleavage assay.

The failure to obtain mutations in the target genes required trouble shooting. Although initial *in vitro* cleavage assays were carried out on two of the three sgRNAs used in the protoplast transformation (Figure 5.13), the results were inconclusive. Addition of the Cas9-sgRNA complex in both cases seemed to reduce the intensity of the band of the target gene, however there was no clear fragment bands present (Figure 5.13). To check the efficiency of the guide RNAs used and potentially develop new and better sgRNAs, the *in vitro* cleavage assays were repeated. Two new sgRNAs each for the target genes RXLR135 and RXLR374 were made, in addition four sgRNAs for a new target gene RXLR325 were designed and ordered (Table 5.4). As well as these, a positive control sgRNA was ordered that had previously been published functioning in *P. capisci* for the target gene ORP1 (Phyca11_564296) (Miao et al., 2018). However it proved challenging to PCR amplify the target gene ORP1, so it was not included in

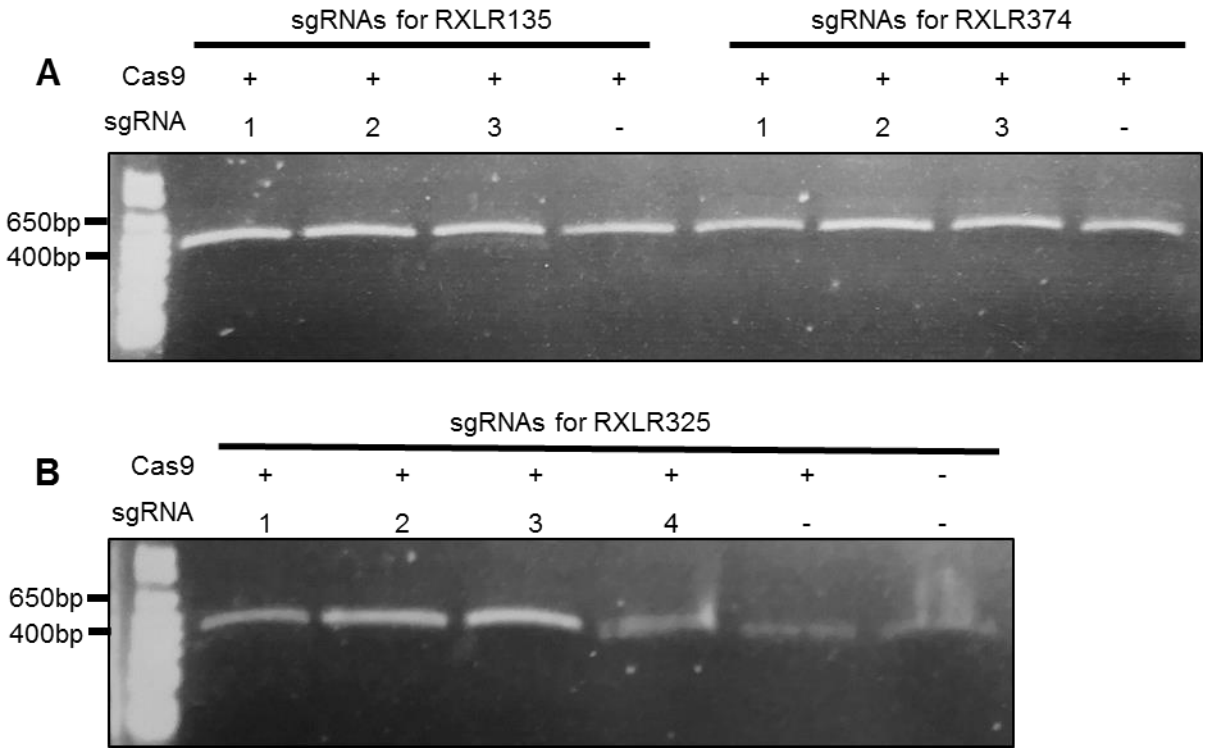


Figure 5.14: *In vitro* cleavage assay of sgRNAs.

Gels showing the digested products of the PCR products of PcRXLR135 and PcRXLR372 and PcRXLR325 following *in vitro* incubation with the Cas9 nuclease and their respective sgRNA., 3 each for PcRXLR135 and PcRXLR372 (A) and 4 for PcRXLR325 (B). Undigested controls with out sgRNA present are also shown. 650 bp and 400 bp DNA band in the ladder are indicated. Expected full length of the PCR product of RXLR 374 is 417bp, of RXLR135 is 386bp, and of RXLR 325 is 423bp . No digested produces can be seen on this gel.

the *in vitro* cleavage assay. *In vitro* cleavage of target genes with these new sgRNA gave no positive results (Figure 5.14). It is unclear why none of these sgRNA were able to cleave their target genes in this experiment.

5.3 Methodology

5.3.1 *P. capsici* culturing

P. capsici strain LT1534 was grown on V8 agar (10% v/v V8 vegetable juice, 1 g/L calcium carbonate, 300 mg/L β-sitosterol, 15% w/v Agar). This strain was used for

both transformation methodologies. *P. capsici* transformants were grown on V8 agar supplemented with section antibiotics, 50 µg/ml G418 and 100 µg/ml Ampicillin. All strains were cultured in the dark at 25°C.

5.3.2 Agro bacterium vector construction

For generation of the pCB301-Cas9-ribo-sgRNA vector, Cas9 was first cloned into the empty pCB301pTOR vector. The full length Cas9 coding gene with an N-terminal nuclear localisation signal (NLS) was amplified from pYF2.3-NLS-Cas9 vector complete with flanking 5' ClaI and 3' EcoRI sites. Both pCB301-TOR and NLS-Cas9 were digested using ClaI (NEB) and EcoRI (NEB) in the 3.1 buffer (NEB). Digested amplicons were ligated into pCB301-TOR using T4 ligase (Promega), generating the pCB301-NLS-Cas9 vector. Next the sgRNA cassette was added to this vector. The full length Ribo-sgRNA cassette was amplified from pYF2.3-ribo-sgRNA plasmid with its 5' RPL41 promotor and its 3' Hsp70 terminator complete with flanking KpnI sites for insertion into pCB301-NLS-Cas9. Both pCB301-NLS-Cas9 and the full ribo-sgRNA cassette were digested using KpnI (Promega) in the buffer J (Promega). Resulting digested fragments were ligated using T4 ligase (Promega), generating the pCB301-Cas9-ribo-sgRNA vector for use with AMT.

5.3.3 *E. coli* and Plasmid preparation

Vectors (pCB301TOR, pYF2-PsNLS-hSpCas9, and pYF515) for storage and amplification were transformed into MACH1 *E. coli* cells (One Shot™ Mach1™ T1 Phage-Resistant Chemically Competent *E. coli*, ThermoFisher). Cells were transformed as per the manufactures instructions. Following transformation, 10-100 µL of cells were plated on pre-warmed (37°C) LB broth agar (containing 10 g/L Bacto-

tryptone, 5 g/L yeast extract, 10 g/L NaCl, 15 g/L agar, and 1 L of dH_2O) supplemented with 100 µg/mL ampicillin plates. Plates were inverted and incubated overnight at 37°C. Positive colonies confirmed by colony PCR were then stored in the form of glycerol stocks and/or used for plasmid amplification via mini or maxi prep. For long term storage, 500 µL of an overnight E. coli culture was added to 500 µL 50% glycerol and kept at -80°C. Plasmids were either collected via maxi (EndoFree Plasmid Maxi Kit) or mini (QIAprep Spin Miniprep Kit, Qiagen) prep depending on downstream use, and as per the manufacturer's instruction.

5.3.4 Agrobacterium transformation

Agrobacterium mediated transformation of Phytophthora was modified from (Wu et al., 2016). Agrobacterium containing the desired plasmid, 2 days prior to transformation were spread onto new LB agar plates, containing the necessary selection antibiotics. To do this one large colony was dissolved in in 50 µL of water and then spread evenly across the entire plate. In 2 days this results in completely carpeted plate. A portion of these cells were then collected by pipetting up a down 5 mL of IM medium. IM medium contains per 1 litre, 800 µL 1.25M K_2HPO_4 pH 4.8, 1 mL 1% $CaCl_2$, 10 mL 0.01% $FeSO_4$, 40 mL 1M MES buffer pH 5.5, 10 mL 50% glycerol, 2.5 mL 20% NH_4NO_3, 20 mL 20% glucose, 5 mL Microelements solution (containing 0.1% w/v of $ZnSO_4.7H_2O$, $MnSO_4.H_2O$, $CuSO_4.5H_2O$, $Na_2MoO_4.7H_2O$, and H_3BO_3), and 20 mL MN buffer (containing 3%w/v $MgSO_4.7H_2O$, 1.5%w/v NaCl) with the addition of 200 µM Acetosyringone). These were the agitated in the dark for 2 hours at room temperature on a shaker at 60 rpm to induce virulence gene expression. After 2 hours on the shaker the optical density (OD) was examined and the agrobacterium suspension was then diluted to 0.4 OD with IM medium.

Zoospores were collected. *P. capsici* LT1534 was grown on 90 mm V8 agar plates in the dark at 25°C, sealed with parafilm. This was followed by 2 days of growth in the light at 22°C without parafilm. Zoospores were collected by flooding the 90 mm plate with 10 mL of sterile distilled water at room temperature, then the mycelia growth was agitated with a sterile plate spreader and everything was transferred to a second *P. capsici* inoculated V8 agar plate. Continuing in this fashion, the 10 mL of water was used to flood and agitate three 90 mm plates recovered into a 50 mL falcon tube. The sporangia suspension was left for 30 minutes with the lid removed on a lightbox to induce zoospore release. The suspension was then filtered through one layer of Miracloth to remove mycelia and any agar chunks in the suspension, this process yielded approximately 5 mL of ≥ 1 x 10^6 spores mL^{-1} zoospore suspension.

Equal volumes of Agrobacterium and zoospore suspension (2-5 mL each) were added together and gently swirled. The mixture was then incubated for 2 hours in the dark at room temperature. Then 500 µL of the mixture was placed on top of a 5 x 5 cm piece of sterile Hybond N+ membrane, which had been placed in a 90 mm solid IM plate, and dried for 10 minutes. IM solid is IM medium as previously described with 1.5% w/v agar. The plates were then incubated in the dark for 2 days at room temperature. After 2 days the membrane was coated in fluffy mycelia, and is then transferred upside down to a Plich medium. Modified Plich media contains per 1 litre, 0.5 g KH_2PO_4, 0.25 g $MgSO_4.7H_2O$, 1 g Asparagine, 1 mg Thiamine, 0.5 g Yeast extract, 10 mg β-sitosterol, 25 g Glucose, and 15 g Agar with 50 µg/mL G418 and 200 µM cefotaxime. It was then incubated in the dark for a further 3 days at room temperature. The membrane was removed and the plated checked at 1-3 days for G418 resistant colonies. These were transferred to V8 medium plates containing 50 µg/mL G418 and 200 µM cefotaxime and checked for expression of transgenes.

5.3.5 sgRNA design and construction

The sequences of the target genes were downloaded from JGI Mycocosm fungal genomic resource, (https://phycocosm.jgi.doe.gov/Phyca11/Phyca11.home.html) LT1534 Assembly v11. Three websites were used to screen for PAM sites and design sgRNAs, the Broad Institute's sgRNA designer, the Eukaryotic Pathogen CRISPR guide RNA/DNA Design Tool (EuPaGDT), and Chopchop. (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design; http://grna.ctegd.uga.edu/; http://chopchop.cbu.uib.no). The top 5 ranked sgRNA were selected from each website, and overlaps discarded. Then any Restriction enzyme sites that sat within the PAM site of the guide RNAs were found using the NEB cutter tool (http://nc2.neb.com/NEBcutter2), and secondary structures were predicted using the 'RNAstructure' web tool (https://rna.urmc.rochester.edu/RNAstructureWeb/). Any sgRNA with free energy secondary structure scores greater than -3.5 were rejected. Resulting sgRNAs were then blasted against the *P. capsici* genome top double check on target effects. (https://fungidb.org/fungidb/showQuestion.do?questionFullName=UniversalQuestions.UnifiedBlast). BLASTN with the E value set to 0.001 and only sgRNA with one hit i.e. the target gene were accepted. The resulting list was manually curated for on-target effects, RNA secondary structure, and overlapping restriction enzymes sites, and the 5 best sgRNA for each gene were chosen. Approved sgRNA sequences were ordered with a 5' HH ribozyme (starting with 6 reverse complement base pairs of the first 6 sgRNA nucleotides) and NheI and BsaI restriction enzyme sites for construction with the pYF515 (Fang et al., 2017) or as described below for IVT and RNP.

5.3.6 Preparation of the pYF515 all-in-one CRISPR-Cas9 plasmid

Preparations of the sgRNAs and pYF515 plasmid were all done according to the (Fang et al., 2017) protocol (Figure 5.7).

5.3.7 sgRNA Oligo Construction

sgRNA constructed were created either for in vitro transcription (IVT) or for insertion into vector for cloning. For construction of oligos for both IVT and vector construction, two complementary oligos were designed with a 20 nt complimentary overlap (Figure 5.7). From 5’ to 3’ the oligos included a T7 promotor sequence, a target specific protospacer and PAM, and the scaffold sequence for *S. pyogenes* Cas9, with the 20nt overlap in the centre. And target specific protospacer portion of the sgRNA sequence had overlapping complementarity in the middle that spanned the guide RNA, with the overlap starting 10 bp in the 3’ direction from the end of the PAM. Note that the sequence has to end in two G’s, which may need to be added if not included at the start of the in target specific protospacer. So if the target sequence starts with one or two G’s the G’s at the end of the T7 promoter with the nucleotides from the protospacer, do not need to be added. PCR was then used to form the full oligo. Oligos were resuspended to 100μM in RNase/DNase free water, and a 100 μL HF Phusion polymerase PCR reaction was set up containing 20 μL of 5x HF buffer, 2 μL of 1 μM top oligo (sgRNA-specific), 2 μL of 1 μM bottom oligo (always the same), 2 μL of 100 μM sgRNA forward primer, 2 μL of 100 μM sgRNA reverse primer, 2 μL of 10 μM dNTPs, 1 μL of HF Phusion polymerase, and 69 μL ddH_2O. The PCR was run in a following manner, in a thermocycler: the sample was held at 98°C for 2 minutes, and then cycle of 98°C for 30 seconds, 60 °C for 30 seconds and finally 72 °C for 30 seconds was repeated 30 times. Following cycling the sample was held at 72 °C for 1

minute and finally was held at 4 °C for maintenance until sample recovery and storage at 4 °C in a refrigerator. To test the successful amplification for sgRNA double stranded DNA template 5 µL of the PCR product was loaded on to a 1% agarose gel dissolved in Tris-Borate-EDTA (TBE) and electrophoresis conducted for 30 minutes at 100 V. TBE buffer contained 45 mM Tris-borate, 1 mM EDTA. Successfully amplified oligo appeared as a 100 bp band. PCR buffer, substrate and enzymes were removed and PCR product purified using QIAquick PCR Purification Kit (Qiagen) as per the manufactures instructions.

5.3.8 *In Vitro* Transcription (IVT)

In order to test the efficiency of sgRNAs, and for the construction of CRISPR ribonucleoprotein (RNP) for both *in vitro* cleavage assays and RNPs for direct RNP mediated mutations, *in vitro* transcription (IVT) of sgRNA from double stranded DNA oligo templates was executed. To conduct IVT Invitrogen™ MEGAshortscript™ T7 Transcription Kit was used. The IVT reaction was assembled in an RNase free microfuge tube, it was composed of a 20 µL reaction as per the manufacturer's instructions. The IVT reaction contained 1 µL of ddH_2O, 8 µL of dNTP Solution mix (75 mM), 8 µL of Template DNA, and 1 µL of T7 Enzyme Mix. The reaction mix was mixed by gentle flicking, and spun briefly in a mini bench top microfuge to collect the reaction in the bottom of the tube. The reaction was incubated at 37°C for 2 hours. Finally, 115 µL nuclease-free water and 15 µL Ammonium Acetate Stop Solution was added. Transcribed RNA was recovered from the solution using acidic Phenol chloroform. 1 volume of Acid-Phenol:Chloroform, pH 4.5 (with IAA, 125:24:1) (Thermo scientific) was added, and the tube was mixed by vortexing. Next, 1 volume of chloroform was added and mixed by inverting 4-6 times. Following a short spin of 30 seconds at 12,000 xg, the aqueous phase was recovered and transferred to a fresh

RNase free microfuge tube. The RNA was precipitated by adding 2 volumes of ethanol and mixing well followed by being chilled for at least more than 30 minutes at -20°C. The precipitated RNA was then pelleted by centrifuging at 4°C for 15 minutes at maximum speed. The supernatant was carefully removed and the RNA pellet was resuspend in RNase free water and stored at –70°C until used.

5.3.9 *In vitro* cleavage assays

The target DNA was amplified using PCR and purified using QIAquick PCR Purification Kit (Qiagen) as per the manufactures instructions. The Molar ration of Cas9 nuclease protein, IVT sgRNA and the target gene was set to 10:10:1, to obtain the best cleavage efficiency. 300 nM sgRNA was prepared by diluting the stock sgRNA from the IVT in RNAse free ice cold water. A dilution of 30 nM of the substrate DNA from the PCR was also prepared in nuclease-free ice cold water. The *in vitro* cleavage reaction was set up, comprising of 20 µL nuclease free water, 3 µL NEBuffer 3.1 (New England Biolabs), 3 µL 300nM sgRNA (for a final concentration 30 nM), 1 µL 1uM Cas9 Nuclease, *S. pyogenes* (New England Biolabs). This mixture was then incubated at for 10 minutes at 25°C followed by the addition of 3 µL of 30 nM substrate DNA. The mixture is pulsed once on a vortex to mix and incubated at 37°C for 30 minutes. The reaction was then stopped by the addition of 1 µL Proteinase K and incubation at room temperature for 10 minutes. To analyse the fragments and success of the reaction, the 30 µL was run on a 1% agarose TBE gel, at 100 v for 40 minutes.

5.3.10 Ribonucleoprotein (RNP) construction

The RNP was constructed either using Cas9 protein for RNP construction (https://international.neb.com/products/m0646-engen-spy-cas9-nls#Product%20Information) or Cas9 protein for sgRNA *in vitro* cleavage assay:

SpCas9 nuclease (New England Biolabs) (https://international.neb.com/products/m0386-cas9-nuclease-s-pyogenes#Product%20Information). The protein was refolded by incubating at 90°C for 5 minutes, followed by cooling to room temperature overnight on the bench. Equimolar, (120 pmol) each, Cas9 protein and sgRNA from IVT were incubated together at 25°C for 10 minutes. RNP was then used immediately for either *in vitro* transcription or for transfection.

5.3.11 Protoplast Transformation

P. capsici LT1534 was cultured anew on V8 solid plate for 2-3 days prior to transformation to encourage new growth. Following this, several agar plugs (5 mm) from the periphery of actively growing cultures were placed into five 200 mL containing nutrient pea broth (NPB) + 100 µg/mL ampicillin. Nutrient Pea Broth (NPB) contains per 1 L 1.0 g K_2HPO_4 (final 0.1%, w/v), 3.0 g KNO_3 (final 0.3%, w/v), 0.5 g $MgSO_4$ (final 0.05%, w/v), 0.1 g $CaCl_2$ (final 0.01%, w/v), 2.0 g $CaCO_3$ (final 0.2%, w/v), 5.0 g d-sorbitol (final 0.5%, w/v), 5.0 g D-mannitol (final 0.5%,w/v), 5.0 g D-glucose (final 0.5%, w/v), 2.0 g yeast extract (final 0.2%, w/v), all made up to 1L with Pea broth, (Pea broth itself is made using 125 g frozen peas placed in 700 mL of water, autoclaved for 30 minutes and then the supernatant collected by straining through four layers of cheesecloth and then brought up to 1 L with ddH_2O), NPB was then autoclaved, and once cool 2 mL vitamin stock (contains 10 µL of 0.02 g/mL Biotin (final 6.7 × 10−7 g/mL), 10 µL of 0.02 g/mL Folic Acid (final 6.7 × 10−7 g/mL), 0.012 g of L-inositol (final 4 × 10−5 g/mL), 0.06 g nicotinic acid (final 2 × 10−4 g/mL), 0.18 g pyridoxine–HCl (final 6 × 10−4 g/mL), 0.015 g Riboflavin (final 5 × 10−5 g/mL), 0.38 g thiamine–HCl (final 1.3 × 10−3 g/mL), H_2O to 300 mL and sterilized using a 0.45 µm filter), and 2 mL trace elements (contains 0.215 g $FeC_6H_5O_7 \cdot 3H_2O$ (final 5. 4 × 10−4 g/mL), 0.15 g

$ZnSO_4 \cdot 7H_2O$ (final 3.8 × 10−4 g/mL), 0.03 g $CuSO_4 \cdot 5H_2O$ (final 7.5 × 10−4 g/mL), 0.015 g $MgSO_4 \cdot H_2O$ (final 3.8 × 10−5 g/mL), 0.01 g H_3BO_3 (final 2.5 × 10−5 g/mL), 0.012 g $Na_2MoO_4 \cdot H2O$ (final 3.0 × 10−5 g/mL), and H_2O to 400 mL then sterilised using a 0.45 µm filter) was added. After 3 days of growth in NPB the mycelia from these flasks were transferred to two larger flasks with 500 mL NPB + 100 µg/mL ampicillin. One big flask provided enough protoplast for 10 transformations. This was allowed to grow for 1.5 to 2 days. At 1.5 days the *P. capsici* mycelia were harvested from the two large flasks using 500 µm filters and, rinsed twice with 0.8 M mannitol. Following washing, all the mycelia was transferred to a 50 mL Falcon tube, and given a further wash with 35 mL 0.8 M mannitol for 10 minutes on a gentle shaker. Whist the Mycelia was washing the enzyme solution was prepared and sterilized using a 0.45 µm filter. The enzyme solution contained 10 mL of 0.8 M mannitol, 8 mL H_2O, 0.8 mL of 0.5 M KCl, 0.8 mL of 0.5 M MES (pH 5.7), 0.4 mL of 0.5 M $CaCl_2$, 0.15 g lysing enzymes from *Trichoderma harzianum*, and 0.06 g Cellulase from *Trichoderma viride*, for a total of 20 mL. Washed mycelia was then harvested using a 500 µm filters and transferred into the prepared enzyme solution and shaken vigorously before left to digest for 40 minutes on a gentle shaker. The digested mycleia, now protoplasts, were then filtered through one layer of miracloth stretched over a large beaker to remove mycelial debris. It is imperative that the protoplast are handled gently throughout. The protoplast solution was then transferred to a new 50 mL falcon tube. This was then centrifuged at 530 xg at 4°C for 4 min. The pellet was then resuspend in 30 mL W5 containing 0.093 g KCl, 4.6 g $CaCl_2 \cdot 2H_2O$, 2.25 g NaCl, 7.8 g glucose, add H_2O to 250 mL, and then spun at 530 xg for 3 min. The pellet was resuspend 10 mL W5 and the protoplasts were left to rest on ice for 30 minutes. They were then spun at 530 xg for 3 min and the W5 removed completely. This time they were resuspended in 5 mL

MMG containing 18.22 g mannitol, 0.76 g $MgCl_2 \cdot 6H2O$, 2 mL of 0.5 M MES (pH 5.7), H_2O made to 250 mL with H_2O. Following suspension in MMG, the concentration of protoplasts was checked using a haemocytometer, and diluted with MMG until the reached approximately 1 million protoplasts per ml. They were again left to rest at room temperature for 10 min. Plasmids were then prepared by adding 20 µg to a fresh 50 mL falcon tube and 1 mL of the protoplast suspension was then added to the plasmids and gently mixed and left on ice for 10 minutes. 1.74 mL of freshly prepared PEG-calcium transformation solution containing 6 g PEG 4000, 3 mL H_2O, 3.75 mL of 0.8 M mannitol, 3 mL of 0.5 M $CaCl_2$ was added in three portions of 580 µL each. The protoplast PEG-calcium solution was gently mixed and incubate on ice for 20 minutes. Following incubation on ice, 2 mL of ice cold PM containing 91.1 g D-mannitol, 1 g $CaCl_2$, 2 g $CaCO_3$, made up to 1 L with Pea broth was added to the tube gently by pipetting down the side. The mixture was then gently mixed by slowly inverting it once and incubated on ice for 2 minutes. 8 mL of ice cold PM was slowly added in the same fashion and again incubated for 2 minutes on ice, then 10 mL ice cold PM was slowly added in the same fashion and again incubated for 2 minutes. Finally, ampicillin was added for a final concentration of 100 µg/mL, and gently mixed as described previously. Protoplasts were then left to allow regenerations, incubated in the dark at 25 °C for 14–18 hours.

The regenerated protoplasts were collected by centrifugation at 700 xg for 5 min at 4°C. The supernatant was poured off and the protoplast pellet was resuspended in 5 mL PM medium. 45 mL warm liquid PAM made from PM with 1.5%, w/v Bacto agar (about 40-50 °C), and containing 30 µg/mL G418 and 100 µg/mL ampicillin, was added to the 5 mL of protoplast. This was inverted twice gently and was then poured approximately equally across three 90 mm petri dishes. After the medium had

solidified the plates were incubated at 25 °C for 2–3 days in the dark. Emerging colonies were covered with a disc of V8 agar medium containing 30 µg/mL G418 and 100 µg/mL ampicillin, and transferred from the petri dish it was formed in using a sterile spatula on top of the PAM – protoplast dish. This was then incubated at 25 °C for 4-7 days in the dark. Throughout this period the plates were checked and emerging viable colonies from transformed protoplast were transferred to a new V8 agar plate containing 30 µg/mL G418 and 100 µg/mL Ampicillin. This was done using a scalpel and excising a sliver of V8 agar containing the top of the colony from the plate. The transformants on V8 agar were incubated for at least 3 days in the dark at 25 °C. Transformants were subsequently tested for expression and mutations.

5.3.12 DNA and RNA extraction

For DNA and RNA isolation of *P. capsici*, agar plugs were transferred to pea broth containing 50 µg/mL G418 and 100 µg/mL Ampicillin and cultured in the dark at 25°C. After 3-5 days of growth, mycelia was separated from the agar plugs and placed into a 2 mL eppendorf tube with a single stainless steel bead and frozen in liquid nitrogen. After disruption using the TissueLyser II (QIAGEN), samples were placed on ice. DNA was isolated using the DNeasy Plant Mini Kit (QIAGEN) as per the manufacturer's instructions. RNA was isolated using the following protocol: 1 mL TRIzol was added to all samples and they were mixed via vortexing until they were completely thawed. 200 µL of chloroform was added and samples shaken vigorously by hand. Samples were centrifuged at 12,000 xg for 10 minutes at 4°C and the upper aqueous phase was transferred into new tubes. 300 µL chloroform was added, samples were mixed by inverting and again centrifuged at 12,000 xg for 10 min at 4°C. Again, 1 volume of isopropanol was added (approx. 500 µL), mixed by inverting and RNA was precipitated

by leaving the samples for 10 min at room temperature or overnight at -20 °C. Samples were then centrifuged for 10 min at 12,000 xg and the supernatant was removed. Pellets were washed twice with 70% ethanol and allowed to air dry for 5-10 minutes. The RNA pellet was then dissolved in 20-50 µL RNAse free water. For cDNA synthesis, RQ1 RNase-Free DNase (Promega) treatment immediately followed by Superscript reverse transcriptase III with Oligo(dT20) was used (Invitrogen) as per the manufacturer's instruction. For sequencing, PCR samples were treated with ExoProStar 1-step enzymatic PCR and sequencing clean-up (Illustra).

5.3.13 Western blot

For total protein isolation from *P. capsici*, agar plugs were transferred to pea broth containing 50 µg/mL G418 and 100 µg/mL ampicillin and cultured in the dark at 25°C. After 3-5 days of growth, isolation mycelia was harvested and lysis buffer GTEN (containing: 10%v/v Glycerol, 25mM Tris buffer pH7.5, 1mM EDTA, 150mM NaCl, 10mM DTT, 2% w/v PVPP, 0.1%v/v tween, and 1mM PMSF) was added. Tissue was disrupted with a TissueLyser II (Qiagen) for 1 minute at 30hz then adapter flip and shake repeated. The resulting lysate was spun at full speed at 4C for 10min. The supernatant was retained and then diluted in equal volumes 2x SDS page buffer, boiled for 5 minutes at 95°C, and loaded onto Mini-PROTEAN TGX precast gel 4-20% (Bio-rad). Separated proteins were then transferred onto a nitrocellulose membrane. The membrane was subsequently blocked with 10% dried skimmed milk (Marvel). Membranes were probed with primary anti-spCas9 (1:1,000) for 1 hour at room temperature. Subsequently, they were incubated with horseradish peroxidase labelled secondary antibodies (1:10,000) (LI-COR) and analysed. The nitrocellulose membrane was then stained with 50% v/v Methanol, 10% v/v Glacial acetic acid, and

40% v/v water containing 0.25% w/v Coomassie Brilliant Blue R-250 for 1 – 2 hours. And, de-stained overnight in water.

5.4 Discussion

The purpose of this study was to increase the genetic tractability of *P. capsici* and develop tools by which more in depth and enlightening gene characterisation studies could be carried out. To that end, the focus has been two fold, 1) to trial transformation techniques for *P. capsici* and 2) to develop the CRISPR Cas9 methodology so it can be utilised in *P. capsici*.

The first transformation technique that was used was agrobacterium mediated transformation, as the methodology had proved successful in *P. palmivora* in implementing the CRISPR/cas9 system (Gumtow et al., 2018). This technique was successful and allowed the creation of transformants, from which transgenes could be amplified from gDNA, and that maintain normal virulence. However there are a few negative aspects to this technique. It requires the creation of bespoke plasmids with T-DNA borders that are necessary for the machinery of Agrobacterium to transfer the genes into the target's genome. It is a fairly labour intensive and complex methodology, and often only produces a few transformants. In addition to these pitfalls, it was found that in the majority of cases, the sgRNA transcripts were not detectable in *P. capsici* transformants. Indeed, whereas protoplast transformation often results in several transgene integration events, AMT usually only results in one or two (Vijn and Govers, 2003), thus affecting possible expression levels. It was decided therefore to utilise a simpler methodology for further experiments with CRISPR Cas9.

Protoplast-PEG transformation was the 2nd transformation technique that was used to attempt to implement the CRISPR Cas9 system in *P. capsici.* Recent publications had

displayed its utility in implementing CRISPR in *P. capsici*, and as such a published plasmid was available (Fang et al., 2017). This plasmid and the new transformation technique together proved to be a more robust transformation method. Although both methodologies are fairly labour intensive and both result in low numbers of transformants, the protoplast transformation method seemed repeatable. Importantly, protoplast transformation generated transformants that expressed the transcripts of the Cas9 and sgRNA, raising the prospect that this approach can be used for gene editing experiments. However, critically and despite the availability of appropriate components of the CRIPSR system, no mutations or deletion events in the target genes were observed in these experiments.

5.4.1 Possible issues in CRISPR/Cas9 in *P. capsici*

Whilst it was expected that it would be possible to implement the CRISPR system in *P. capsici,* as other groups have shown (Wang et al., 2019, Miao et al., 2018), as well as in the other *Phytophthora* species (Fang and Tyler, 2016, Pettongkhao et al., 2020), implementing the system proved challenging. In fact other species of *Phytophthora* species appear recalcitrant to CRISPR gene editing, despite exhaustive testing (van den Hoogen and Govers, 2018). There are a number of possible reasons that these experiments, using the CRISPR system described, were unable to produce mutations in the target genes. While some of these issues may be resolved by troubleshooting exercises, others may mean that the CRISPR system has no utility in *P. capsici*. However, this study has been by no means exhaustive in our testing of the CRISPR system in *P. capsici*, and indeed further work is required to evaluate and draw conclusions on the ability and utility of CRISPR based gene editing in *P. capsici*

One possible reason that no mutations occurred could be the low or no efficacy in the sgRNA, and therefore inefficient targeting of the target genes. There is a test for sgRNA efficiency beyond the computer predicted sgRNA generating software that was utilised in the sgRNAs design, this being *in vitro* cleavage assays. These can be used to test the efficiency of the sgRNA before transformation, however due to the lack of any good positive controls, and inconclusive results it is unclear if the sgRNAs that were designed for this study were able to cleave their target genes with sufficient efficacy and specificity.

However, there is also some evidence, suggesting that the Cas9 protein is truncated or being degraded in *P. capsici*. The presence of a full size Cas9 protein by Western blot was never detected in these experiments. Only bands corresponding to truncated or degraded Cas9 protein could be detected in the transformants. It is unclear whether these represent degradation products of the Cas9 protein, and that if they do, whether this is the cause of the lack of Cas9 mediated mutations. It could be the case that even if degraded the full length Cas9 in complex with the sgRNA, prior to any degradation event could still cause a mutation in the target gene, or it could be that the Cas9-sgRNA complex never even reaches the nucleus (even with the presence of the NLS from *P. sojae*). Issues that relate to the integrity and localisation of the Cas9-sgRNA complex would be hard to troubleshoot in comparison to designing and identifying sgRNAswith good efficiency. Other versions of Cas9 i.e. not derived from *Streptococcus pyogenes,* or more *P. capsici* specific nuclear localisation signals may need to be constructed.

5.4.2 Future work

Initially, further work requires the troubleshooting of the *in vitro* cleavage assay to confirm the efficacy of the designed sgRNAs. Once the effectiveness of these sgRNAs has been confirmed, new sgRNAs have designed and validated *in vitro*, other areas can be explored to test the CRISPR/Cas9 system in *P. capsici*. Thus, moving forward there are two main areas for exploration of gene editing in *P. capsici*. For the purpose of the study, the aim of this work was to quickly and completely knock out key genes that were thought to have a role in dynamic host adaptation. For this, expression of a single sgRNA, alongside the Cas9 protein, was required. One methodology to do this, is the delivery of ribonucleoproteins (RNP) directly into *P. capsici*. This approach involves the construction of the Cas9-sgRNA complex *in vitro*, followed by incubation of these complexes with *P. capsici* protoplasts in a manner very similar to the protoplast transformation methodology. The transient exposure to readymade guided Cas9 complexes removes some of the possible issues with the integration and expression of Cas9 and sgRNA, and their ability to form a complex in the *P. capsici* cell. Although there is still the issue of nuclear localisation and sgRNA effectivity, the approach may work in *P. capsici*. It should be noted however, this methodology does not appear to work in *P. infestans*, (van den Hoogen and Govers, 2018). Testing a wider range of methodologies across systems will help untangle the technical and biological reasons for failure.

Other studies have successfully implemented the CRISPR system in *P. capsici* (Wang et al., 2019, Wang et al., 2018). The largest difference between the attempt made here and those already published is the fact that the aim of those studies was to introduce specific point mutation in a target gene (in this case PcORP1). Thus the methodology required two separate sgRNAs targeting the same gene, and a repair template with the desired point mutation that sits between and extends beyond either side of the two

sgRNA cut sites. The repair template is incorporated into the target gene by the cell's own repair mechanisms, thus incorporating the desired mutation. The presence of two sgRNAs and a repair template may be desirable when conducting CRISPR experiments in *P. capsici*. Additionally, if used simply to knock out a gene the repair template can be used to introduce a selection gene which would aid the selection of successful transformants and knock outs. Although more laborious and complex than a single sgRNA, the introduction of a selective repair template may be necessary for the functioning of the CRISPR system in *P. capsici*. It would also require the successful transformation of multiple plasmids instead of one which may limit the functionality.

5.4.3 Conclusion

Although this study has been unsuccessful in introducing the CRISPR/Cas9 into *P. capsici* for the knockout of target genes. Two methodologies for the transformation of *P. capsici* were trialled and successfully implemented. This has increased the molecular and genetic malleability of *P. capsici*. This study has also made the path forward clearer towards the aim of implementing CRISPR/Cas9 in *P. capsici* in the future.

Table 5.2: Table of oligos used in this study.

Application	Oligo ID	Sequence	Target
Transformant screening	M13PF (-20)	GTAAAACGACGGCCAG	pYF515 and pCB301TOR-PsNLS-hSpCas9-ribo-sgRNA
	M13PR	CAGGAAACAGCTATCAC	
	KAN_forwarda	GATGGATTGCACGCAGGTTC	
	KAN_reversea	GATGGATTGCACGCAGGTTC	
	GRNAcassette_forwarda	GATGACGTGTGCTCTTCCACC	
	GRNAcassette_reversea	CAGAACACCCATACACGCTG	
	PsCas9_forwarda	GATACACCAGACGGAAGAACCG	
	PsCas9_reversea	TGGTCATCCAGGCGAATCTGC	
	Cas9_JQ_FP	AGTCCGGCAAGACAATCCTG	
	Cas9_JQ_RP	TTCCGGAAATCGGACACCAG	
	gRNA_JQ_FP	CATGCAAACCAGCCATAGTG	
	gRNA_JQ_RP	AACATCGTTCGGGGTAAGTG	
Sequencing primers	P_534478_FP	TGCGTCTGTCCTTCCTTCTG	RxLR135
	P_534478_RP	CGAAATAAGCCCAGGAGCGA	
	P_13944_FP	GTCTCTCCTACGCCCTCCC	RxLR373
	P_13944_RP	CAAGTAATCACGGTACTTCTGACG	
	P_13953_FP	ATGCGTCTTTCCTACACTCTCC	RxLR374
	P_13953_RP	TACAAGAAATCACGGTACTTCTGAC	
	P_21535_FP	GCTGCCTTCACATTGCTCTT	RxLR001
	P_21535_RP	TCTCCCCCAGACCATGTAGG	
	PcRxLR325_FP	GCGTCTCACTTACATTCTCGC	RxLR325
	PcRxLR325_RP	TGACCATTTCTTTCGCTCTTGC	
pCB301TOR-PsNLS-hSpCas9-ribo-sgRNA construction	504243_ForClaI_Flag	AGTCATCGATACCATGGACTACAAAGACGATGAC	2nd amplification of GOI N terminus w/ ClaI
	504243_ForFlag_Gene	GACTACAAAGACGATGACGACAAGGCTGCCACCTTTGTGCCGAC	1st amplification of GOI N terminus w/ Flag tag
	504243_RevGene_EcoRI	CAGTCGAATTCCTAATCAAACGTCACGAGGTG	1st/2nd amplification of GOI C terminus w/ EcoRI
	565968_ForClaI_Flag	AGTCATCGATACCATGGACTACAAAGACGATGAC	2nd amplification of GOI N terminus w/ ClaI
	565968_ForFlag_Gene	GACTACAAAGACGATGACGACAAGCAGGACATTCAGATCATG	1st amplificaton of GOI N terminus w/ Flag tag
	565968_RevGene_EcoRI	CAGTCGAATTCAATCTAGCAGTTGGG	1st/2nd amplification of GOI C terminus w/ EcoRI
	107224_ForClaI_Flag	AGTCATCGATACCATGGACTACAAAGACGATGAC	2nd amplification of GOI N terminus w/ ClaI r
	107224_ForFlag_Gene	GACTACAAAGACGATGACGACAAGGTCGAGGCGTACACGATA	1st amplification of GOI N terminus w/ Flag tag
	107224_RevGene_EcoRI	CAGTCGAATTCTTACAAAGGGTAATGTCCACCG	1st/2nd amplification of GOI C terminus w/ EcoRI
	EcoR1 cas9 rev	CAGTCGAATTCTTAGTCGCCTCCCAGCTGAG	Amplification of Cas9 from C terminus w/ EcoRI
	For_pYF2.3G-sgRNA-KpnI	TCGGTACCCCCCAATTCCCCGGATCG	Amplification of the sgRNA cassette from N terminus w/ KpnI
	Rev_pYF2.3G-sgRNA-KpnI	TCGGTACCGTTCCGTCATTTCCTCGCAGC	Amplification of the sgRNA cassett from C terminus w/ KpnI
	For_pYF2.3G-cas9-ClaI	AGTCATCGATACCATGCACAAGCGCAAGCGC	Amplification of Cas9 from N terminus w/ ClaI
	Rev_pYF2.3G-cas9-EcoRI	CAGTCCTTAAGTTAGTCGCCTCCCAGCTGAG	Amplification of Cas9 from C terminus w/ EcoRI

Table 5.3: Table of sgRNAs used in the transformation.

The whole oligos for sgRNA cassette structure are no shown here just the PAM and Protospacer sections.

sgRNAs	668 fwd	GATATCGTCATACAGCAGCTCGG	ABC1
	JQ2_148_revcom	TCACTGGAAAGGACTTGGTTTGG	RxLR135
	JQ4_86_revcom	AGCAATTATGAGGAGAGTGTAGG	RxLR374
	JQ4_141	CCACCCGTTGACTTTCTTCCAGG	RxLR001

Table 5.4: sgRNA used in in vitro cleavage assay.

Target sequence	Orientation	PAM site	Restriction enzyme	2ndry structure energy	BLAST Offtarget
PHYCA_534478 (PcRxLR135)					
GCCACTTGACAGGAGCG CCACGG	anti	63	BtgI, BSaJI	-2.4	no
CCACGGCAGCAACCAACA GAAGG	anti	46	BsII, HpyAV	-2.8	no
CTGGGTCACTAAGCGCTA CACGG	sense	294	BciVI	-0.2	no
PHYCA_13944 (PcRxLR373)					
CGTCGTCGCTCTCATCCT CGGGG	anti	97	BsaJI, BsoBI	-1.6	no
CGACGACGAAGCGTTGCT TGAGG	sense	130	MnlI	-0.7	Yes
AACAGCGTTTGGCGACGT CATGG	anti	16	FatI, CviAII, NlaIII	-3.1	Yes
PHYCA_13953 (PcRxLR374)					
CACAACCAACGGTCGCAC CACGG	sense	82	BtgI, BsaJI,	-1.3	no
CCACAACAACGAGGGCCT CGAGG	sense	184	SmlI, AvaI, BsoBI, PaeR71, PspXI, XhoI, MnlI	-2.6	no
TCTGCGTTACCACAACAA CGAGG	sense	175	MnlI	-2.5	no
PHYCAscaffold_50_f_194674_195096_1 (PcRxLR325)					
ACTTCTTCAAAACATCTCC GAGG	anti	266	BsaJI, Hpy188I, MnlI	-0.1	no
CATCGAGGTCATCGTAAT CGAGG	anti	206	TaqI, MnlI	-2.1	no
CTCATTGCTGATGACGAC GCCGG	sense	170	BsaHI, BsrFI, MspI, HpaII,	-1.2	no

Chapter 6: General Discussion

6.1.1 The aims, opportunities and problems

The majority of studies that investigate the molecular dynamics of plant pathogen interaction are done on single pathogen-hosts systems (Dong et al., 2015). Investigations of broad host ranged pathogens and the divergence in the molecular mechanics of dynamic host adaption have given insight into pathogen biology (Mbengue et al., 2016). Much of this research has been done on the genomic mechanisms by which new hosts can be incorporated into the host range, “hosts jumps” (Ma et al., 2010, Kellner et al., 2014, Raffaele et al., 2010). A few studies have profiled the gene expression of a pathogen on multiple host plants revealing suites of differentially expressed genes that may have a role in the dynamic adaption of a pathogen to a particular host plant (Yang et al., 2018, Allan et al., 2019, Harris et al., 2016). This expands our understanding of how broad host ranged pathogens can maintain virulence of multiple host plants and reveals the key molecular factors in host specificity and host range.

Preliminary data acquired by the Huitema lab that was explored in Chapter 1, showed that differential gene expression may be the key to dynamic host adaptation of the board ranged oomycete pathogen *P. capsici* to distinct host plants. The aim of this project, therefore, was to take an unbiased and broad approach to examining the molecular factors in the host specific dynamic adaption of *P. capsici*. As such, broad “omics” style examination of the effect that distinct host had on *P. capsici* was deemed to be the best approach. However, this approach introduces problems when it comes to examining the infectious stages of the pathogen life cycle. As has been discussed in several chapters, it is problematic to isolate *P. capsici* tissue from infected plant

tissue, especially at the early stages of infection where the biomass of pathogen to plant is very low. Thus, one of the key aims of this project was to develop tools to enable the ability to either circumvent or remove this issue of excessive plant material contaminating *P. capsici* samples.

6.1.2 The solution

An elegant solution that was originally trialled, was what is termed translating ribosomal affinity purification followed by RNA sequencing (TRAP-SEQ). However multiple attempts to optimise this methodology to allow the isolation of *P. capsici* translating mRNA from infected leaf tissue met with little to no success. The root of the issue was never resolved; however it could be suggested that either limited biomass of the pathogen in infected leaf tissue, or some chemical element of the host leaves used meant that the isolation of mRNA from purified ribosomal protein was not possible in this setting. Although, it was possible to isolate translating mRNA from *P. capsici* mycelia, which could be used in future experiments to analyse the translating RNAs of the non-infectious lifecycles stages of *P. capsici*.

Following this, a simpler methodology that involved *in vivo* inoculation of host leaf extracts was devised. Although, it was assumed that this would represent less resemblance to the infection of plants, validation experiments showed similar differential gene expression to that shown in Chapter 1, data gained from Microarray analysis of *P. capsici* gene expression in 4 host plants. Based on these results two omics style experiments using this extracts inoculation technique were conducted.

6.1.3 The findings and context

From these experiments, insights into the molecular characteristics of dynamic adaption have been gained. For the purposes of comparisons between hosts, two host plants from two different taxonomic families that the Microarray experiments in Chapter 1 had shown to be most different, were used, tomato and cucumber. Both transcriptomics and proteomics were conducted on *P. capsici* inoculated in these host extracts (TE and CE). It had been hypothesised that the key event, especially the regulation events of dynamic host adaption would take place in the early stages of infection, therefore sample from *P. capsici* inoculated TE and CE were taken at 2, 4 and 8 hpi. For these experiments GC at 2hpi were used as a control.

Both proteomics and transcriptomics studies found that inoculation in host extract had an effect on the gene/protein expression, it was also found that the two different hosts had had a different effect when it came to protein intensity or transcript expression, on *P. capsici.* Although the specific genes that both studies highlighted were not similar, the broad patterns of expression remained the same. It was shown that challenging *P. capsici* with TE induced fewer genes than challenge with CE, but what gene/ proteins that are differentially up-regulated in TE happen at early time points. Those proteins that were found with increased intensity in the early time points in TE seemed to have a role in regulation and structural changes. Another of the main finding, that is backed up by the literature, shows that a great deal of DE genes/proteins involved in oxidoreductase activity, in both host extracts (Srivastava et al., 2013, Kellner et al., 2014, Ma et al., 2010). And it has been suggested that these are to deal with innate toxins and ROS found in host tissues (Yang et al., 2018, Pang et al., 2016, Giannakopoulou et al., 2014). It was also found that transporter and membrane proteins play a key role in dynamic host adaption. Proteomics also revealed that there

is a dramatic metabolic and respiratory shift that occurs in GC, similar to previous phenotypes of *Phytophthora* cyst germination reported in the literature (Grenville-Briggs et al., 2005, Pang et al., 2017, Hosseini et al., 2015). Interestingly though we found a similar metabolic and respiratory focused protein shift in the CE, this was not found in TE. These gene families may be of interest to explore as determinates of host range, however whether these genes are host specific or just a broader mechanism of necessary adaption to the host environment is unclear. Although oxidoreductase proteins have been proposed a possible host specificity factors (Li et al., 2020) and frequently found upregulate in omics experiments in pathogen – host interactions (Srivastava et al., 2013, Kellner et al., 2014, Ma et al., 2010) in this study they were also found upregulated in many GC samples, hinting and their utility in spore germination.

The study, especially the data received from the proteomics study in Chapter 4 also had utility in revealing the mechanisms of cyst germination in water. Much of what was observed hinted towards a starvation phenotype. Changes in metabolism, proteasomal subunits, stress response proteins, and changes in amino acid synthesis all hint as a starvation phenotype, and indeed other studies of cyst germination hint at a similar phenotype (Resjö et al., 2017, Grenville-Briggs et al., 2005). This suggests that a single zoospore upon contact with a host has to have the ability to bootstrap the infection and perhaps the initial stages of colonisation and ETS, prior to its ability to receive nutrients from the host plant. In both Chapters 3 and 4, it was shown that CE had a marked similarity to the GC samples, a similarity not found in TE, and that TE was marked by a down regulation event. It could be suggested therefore that if the difference between the two hosts and GC is one of access to nutrients, and that *P.*

capsici is, therefore, better able, perhaps adapted to receive nutrients from TE than CE, explaining the potential partial starvation phenotype shown in CE.

Further evidence to support the alternative hypothesis that host dependant difference in gene and protein expression found in *P. capsici* cysts germinating in CE and TE is due to the difference in nutrient availability, is the lack of substantial gene upregulation in TE samples. Although the distinction here is perhaps more between direct host perception and a host perception model based on nutrient availability. It has been stated in previous chapters that in the lab the constantly grown on media largely derived from tomato fruit juice. Previous studies have evaluated the mechanism of epigenetic changes in *Phytophthora* species (Whisson et al., 2012, Chen et al., 2018, van West et al., 2008) and have shown their effectiveness in effector expression (Tzelepis et al., 2020, Kasuga and Gijzen, 2013). Epigenetic adaptation from long term growth on a tomato based media could account for the lack of response to inoculation in TE relative to the response to CE and GC. Some studies suggest that epigenetics, or perhaps some other mechanism could represent a more medium term mechanism for adaptation environment and indeed to distinct host plants (Kasuga and Gijzen, 2013). Sudden epigenetic modifications in *Phytophthora* species have been shown to accompany the adaptation of a pathogen to a novel environment or presentation with novel host plant (Bossdorf et al., 2008), such movement and novel host presentation is commonplace in a global agriculture system (Kasuga and Gijzen, 2013).

In essence, however both these hypotheses, whether extract can induce differences in expression by direct perception of host elements, or by indirect perception of the nutrients available in the environment, host perception by the pathogen is taking place. An alternative hypothesis would surround the adaptability of large very heterogeneous populations, that upon confrontation to with the new host environment those

zoospores better adapted to the new environment survive. Dynamic adaption via natural selection of a broad heterogeneous population (Gilbert and Whitlock, 2017). However, despite the apparent logic of this hypothesis, this was not tested in this study. Although the longer-term shift in population makeup has been found to represent evolutionary adaptation in plant pathogens (Susi et al., 2020) what effect this kind of shift has in dynamic short term adaptation is unclear, and perhaps understudied.

The second main aim of this project was to develop tools for the investigation of *P. capsici* to aid omic style investigation and increase the genetic tractability. It was shown that TRAP-SEQ as it is at the moment is not a viable technology for the investigation of *P. capsici* translateome during leaf infection, so the alternative *in vivo* extract inoculation assay was developed to allow transcriptomic and proteomic studies. As discussed here and in Chapters 3 and 4 these identified many genes of interest that may have a role in host specific virulence and dynamic adaption. To aid in the characterisation of these genes therefore this project attempted to develop the CRISPR/Cas9 gene editing technology. This was to be used to quickly and efficiently totally knockdown genes of interest. During the development of the CRISPR system in *P. capsici* two separate transformation methods (AMT, and PEG/$CaCl_2$) were optimized and used to insert the CRISPR system transgenes, although these methodologies were successful, the CRISPR system seemed to not be functional in *P. capsici.* This is despite publications stating the opposite (Wang et al., 2019, Miao et al., 2018). here it was found that no mutation occurred even with the expression of sgRNA and Cas9. One potential problem was the lack of a full size Cas9 protein identifiable by western blot, instead multiple smaller fragments were identified. No data was collected by there seemed to be no growth deficiencies in those strains

expressing Cas9, however recently it was suggested that the active nuclease is toxic to *P. infestans* (Ah-Fong et al., 2021). They were able to utilise the Cas12a nuclease as a replacement for Cas9 in the CRISPR system in order to induce mutations in P*. infestans*, presumably Cas12a is not toxic to *P. infestans* and perhaps other *Phytophthora* species {Ah-Fong, 2021 #255}. The utility of the CRISPR system in *Phytophthora* species is still an open question, and implementation in certain strains and species may require considerable troubleshooting. Going forward, further optimisation of the CRISPR/system in *P. capsici* would likely enable gene knockout and gene editing; whether this is simply the design of new sgRNAs, the use of dual sgRNAs and a repair template, or the use of the Cas12a protein as an alternative, remains to be determined. Other methods of delivery such as RNP could also be trialled.

This would allow for the characterisation of the genes identified in the experiments of Chapters 3 and 4. Particularly the investigation of key differentially expressed oxidoreductase activity genes, transporter genes (especially ABC transporter). The RAP-GTPAses that were found uniquely induced in TE would also be worth investigating in this fashion. We also found a series of effectors (RXLRs and CRN) expressed differentially in the host extracts which would be of interested when investigating host specific virulence strategies.

6.1.4 Conclusion and Further work

As stated, this project had two overarching aims, 1) develop tools for the examination of *P. capsici*, and 2) to investigate host specific virulence and dynamic adaption. In both these aims this project was partially successful. Two transformation methods where optimised. In addition a methodology that allows the isolation of translating

mRNA from non-infectious stages of *P. capsici* was developed. Head way was also made into optimising the CRISPR/Cas9 system although more work in this area is needed.

Insights into potential mechanisms of dynamic adaption were gained. This project has highlighted many exciting new angles, and specific gene and proteins for future studies into host specific virulence and dynamic adaption. Whether the differences in host gene/protein expression are due to a direct host perception event, due to the differences in available nutrients, temporal differences due to the suitability of the environment, or some effect of selection on a heterogeneous population is still unclear. However all of these represent a system by which *P. capsici* can dynamically adapt to a host. A more open question is whether the difference in *P. capsici* response to the host extracts represents an actual adaptive mechanism that would also be observed in leaf infection, or whether confrontation of zoospores with the internal environment of the leaf, and leaf cells induces a natural response or a response that is an artefact of inoculation in a plant extract. That being GC would represent a starvation state, TE would represent adequate nutrients and a vegetative state, and CE perhaps some middle ground between starvation and adequate nutrients availability. Expansion of the extract experiment to include a known nutrient broth that is used for *P. capsici* lab growth such as Pea broth, or V8 vegetable juice, as well as the inclusion of extracts from non-plants may be of interest to try to parse out the effect nutrient availability has on cyst germination.

In addition to further investigation of the effects of nutrient and host extracts on germination cysts, this study has identified a large suite of genes, and gene families that could be involved in dynamic adaption to the host environment. Investigation of these genes, especially those from transporter families and with oxidoreductase

activity will enable analysis of the importance of these genes to host virulence and adaption and thus may enable further understanding of the mechanisms that underlie dynamic adaptation. Especially, if these genes as hypothesised, and shown in the literature (Morris and Phuntumart, 2009, Judelson and Senthil, 2006, Srivastava et al., 2013) enable colonisation of an otherwise toxic host environment. This work has increased our understanding and suggested new possible routes to investigate the mechanism of broad host range in devastating crop pathogens.

References

AH-FONG, A. M., BORMANN-CHUNG, C. A. & JUDELSON, H. S. 2008. Optimization of transgene-mediated silencing in Phytophthora infestans and its association with small-interfering RNAs. *Fungal Genet Biol,* 45**,** 1197-205.

AH-FONG, A. M. V., BOYD, A. M., MATSON, M. E. H. & JUDELSON, H. S. 2021. A Cas12a-based gene editing system for Phytophthora infestans reveals monoallelic expression of an elicitor. 22**,** 737-752.

AH-FONG, A. M. V., KIM, K. S. & JUDELSON, H. S. 2017. RNA-seq of life stages of the oomycete Phytophthora infestans reveals dynamic changes in metabolic, signal transduction, and pathogenesis genes and a major role for calcium signaling in development. *BMC Genomics,* 18**,** 198.

ALLAN, J., REGMI, R., DENTON-GILES, M., KAMPHUIS, L. G. & DERBYSHIRE, M. C. 2019. The host generalist phytopathogenic fungus Sclerotinia sclerotiorum differentially expresses multiple metabolic enzymes on two different plant hosts. *Sci Rep,* 9**,** 19966.

AMARO, T. M., THILLIEZ, G. J., MOTION, G. B. & HUITEMA, E. 2017. A Perspective on CRN Proteins in the Genomics Age: Evolution, Classification, Delivery and Function Revisited. *Front Plant Sci,* 8**,** 99.

ANDRIVON, D., PILET, F., MONTARRY, J., HAFIDI, M., CORBIERE, R., ACHBANI EL, H., PELLE, R. & ELLISSECHE, D. 2007. Adaptation of Phytophthora infestans to Partial Resistance in Potato: Evidence from French and Moroccan Populations. *Phytopathology,* 97**,** 338-43.

ANTONOVICS, J., BOOTS, M., EBERT, D., KOSKELLA, B., POSS, M. & SADD, B. M. 2013. The origin of specificity by means of natural selection: evolved and nonhost resistance in host-pathogen interactions. *Evolution,* 67**,** 1-9.

APINYA, K., OYTHIP, P., PANITI, P., AKKAWAT, T., SOPONE, W. & PIYADA, A. T. 2018. Effects of Culture Media on Phytophthora palmivora Growth, α-elicitin Production and Toxicity to Dendrobium. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca,* 46.

BABADOOST , M., TIAN, D., ISLAM, S. Z. & PAVON, C. 2008. Challenges and options in managing Phytophthora blight (Phytophthora

capsici) of cucurbits1. *Proceedings of the IXth EUCARPIA meeting*.

BAHRI, B., KALTZ, O., LECONTE, M., DE VALLAVIEILLE-POPE, C. & ENJALBERT, J. 2009. Tracking costs of virulence in natural populations of the wheat pathogen, Puccinia striiformis f.sp.tritici. *BMC Evol Biol,* 9**,** 26.

BALL, B., LANGILLE, M. & GEDDES-MCALISTER, J. 2020. Fun(gi)omics: Advanced and Diverse Technologies to Explore Emerging Fungal Pathogens and Define Mechanisms of Antifungal Resistance. *mBio,* 11**,** e01020-20.

BARCHENGER, D. W., LAMOUR, K. H. & BOSLAND, P. W. 2018. Challenges and Strategies for Breeding Resistance in Capsicum annuum to the Multifarious Pathogen, Phytophthora capsici. *Frontiers in Plant Science* [Online], 9. Available: http://europepmc.org/abstract/MED/29868083

https://doi.org/10.3389/fpls.2018.00628

https://europepmc.org/articles/PMC5962783

https://europepmc.org/articles/PMC5962783?pdf=render [Accessed 2018].

BERTIN, B., RENAUD, Y., ARADHYA, R., JAGLA, K. & JUNION, G. 2015. TRAP-rc, Translating Ribosome Affinity Purification from Rare Cell Populations of Drosophila Embryos. *J Vis Exp*.

BIRCH, P. R. J. & WHISSON, S. C. 2001. Phytophthora infestans enters the genomics era. *Molecular Plant Pathology,* 2**,** 257-263.

BOS, J. I., ARMSTRONG, M. R., GILROY, E. M., BOEVINK, P. C., HEIN, I., TAYLOR, R. M., ZHENDONG, T., ENGELHARDT, S., VETUKURI, R. R., HARROWER, B., DIXELIUS, C., BRYAN, G., SADANANDOM, A., WHISSON, S. C., KAMOUN, S. & BIRCH, P. R. 2010. Phytophthora infestans effector AVR3a is essential for virulence and manipulates plant immunity by stabilizing host E3 ligase CMPG1. *Proc Natl Acad Sci U S A,* 107**,** 9909-14.

BOS, J. I., KANNEGANTI, T. D., YOUNG, C., CAKIR, C., HUITEMA, E., WIN, J., ARMSTRONG, M. R., BIRCH, P. R. & KAMOUN, S. 2006. The C-terminal half of Phytophthora infestans RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in Nicotiana benthamiana. *Plant J,* 48**,** 165-76.

BOSSDORF, O., RICHARDS, C. L. & PIGLIUCCI, M. 2008. Epigenetics for ecologists. *Ecol Lett,* 11**,** 106-15.

BOWYER, P., CLARKE, B. R., LUNNESS, P., DANIELS, M. J. & OSBOURN, A. E. 1995. Host range of a plant pathogenic fungus determined by a saponin detoxifying enzyme. *Science,* 267**,** 371-4.

BOZKURT, T. O. & KAMOUN, S. 2020. The plant–pathogen haustorial interface at a glance. *Journal of Cell Science,* 133.

CAI, R., YAN, S., LIU, H., LEMAN, S. & VINATZER, B. A. 2011. Reconstructing host range evolution of bacterial plant pathogens using Pseudomonas syringae pv. tomato and its close relatives as a model. *Infect Genet Evol,* 11**,** 1738-51.

CHEN, H., SHU, H., WANG, L., ZHANG, F., LI, X., OCHOLA, S. O., MAO, F., MA, H., YE, W., GU, T., JIANG, L., WU, Y., WANG, Y., KAMOUN, S. & DONG, S. 2018. Phytophthora methylomes are modulated by 6mA methyltransferases and associated with adaptive genome regions. *Genome Biology,* 19**,** 181.

CHEN, X., SHEN, G., WANG, Y., ZHENG, X. & WANG, Y. 2007. Identification of Phytophthora sojae genes upregulated during the early stage of soybean infection. *FEMS Microbiol Lett,* 269**,** 280-8.

COLEMAN, J. J., ROUNSLEY, S. D., RODRIGUEZ-CARRES, M., KUO, A., WASMANN, C. C., GRIMWOOD, J., SCHMUTZ, J., TAGA, M., WHITE, G. J., ZHOU, S., SCHWARTZ, D. C., FREITAG, M., MA, L. J., DANCHIN, E. G., HENRISSAT, B., COUTINHO, P. M., NELSON, D. R., STRANEY, D., NAPOLI, C. A., BARKER, B. M., GRIBSKOV, M., REP, M., KROKEN, S., MOLNAR, I., RENSING, C., KENNELL, J. C., ZAMORA, J., FARMAN, M. L., SELKER, E. U., SALAMOV, A., SHAPIRO, H., PANGILINAN, J., LINDQUIST, E., LAMERS, C., GRIGORIEV, I. V., GEISER, D. M., COVERT, S. F., TEMPORINI, E. & VANETTEN, H. D. 2009. The genome of Nectria haematococca: contribution of supernumerary chromosomes to gene expansion. *PLoS Genet,* 5**,** e1000618.

COX, J., MATIC, I., HILGER, M., NAGARAJ, N., SELBACH, M., OLSEN, J. V. & MANN, M. 2009. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc,* 4**,** 698-705.

DE SOUSA ABREU R, P. L., MARCOTTE EM, VOGEL C. 2009. Global signatures of protein and mRNA expression levels. . *Mol Biosyst.,* 5**,** 1512-26.

DE VIENNE, D. M., REFREGIER, G., LOPEZ-VILLAVICENCIO, M., TELLIER, A., HOOD, M. E. & GIRAUD, T. 2013. Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. *New Phytol,* 198**,** 347-85.

DEAN, R., VAN KAN, J. A., PRETORIUS, Z. A., HAMMOND-KOSACK, K. E., DI PIETRO, A., SPANU, P. D., RUDD, J. J., DICKMAN, M., KAHMANN, R., ELLIS, J. & FOSTER, G. D. 2012. The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol,* 13**,** 414-30.

DELAUNOIS, B., JEANDET, P., CLÉMENT, C., BAILLIEUL, F., DOREY, S. & CORDELIER, S. 2014. Uncovering plant-pathogen crosstalk through apoplastic proteomic studies. *Frontiers in Plant Science,* 5**,** 249.

DEPOTTER, J. R. L., SEIDL, M. F., WOOD, T. A. & THOMMA, B. P. H. J. 2016. Interspecific hybridization impacts host range and pathogenicity of filamentous microbes. *Current Opinion in Microbiology,* 32**,** 7-13.

DIE, J. V. & ROMÁN, B. 2012. RNA quality assessment: a view from plant qPCR studies. *Journal of Experimental Botany,* 63**,** 6069-6077.

DODDS, P. N. & RATHJEN, J. P. 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat Rev Genet,* 11**,** 539-48.

DONG, S., RAFFAELE, S. & KAMOUN, S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. *Curr Opin Genet Dev,* 35**,** 57-65.

DONG, S., STAM, R., CANO, L. M., SONG, J., SKLENAR, J., YOSHIDA, K., BOZKURT, T. O., OLIVA, R., LIU, Z., TIAN, M., WIN, J., BANFIELD, M. J., JONES, A. M., VAN DER HOORN, R. A. & KAMOUN, S. 2014. Effector specialization in a lineage of the Irish potato famine pathogen. *Science,* 343**,** 552-5.

DRENTH, A. & GUEST, D. I. 2004. Diversity and management of Phytophthora in Southeast Asia. *Australian Centre for International Agricultural Research (ACIAR)*.

ERWIN, D. C. & RIBEIRO, O. K. 1996. Phytophthora Diseases Worldwide. *American Phytopathological Society, St. Paul, MN.*

ESCOBAR-NIÑO, A., LIÑEIRO, E., AMIL, F., CARRASCO, R., CHIVA, C., FUENTES, C., BLANCO-ULATE, B., CANTORAL FERNÁNDEZ, J. M., SABIDÓ, E. & FERNÁNDEZ-ACERO, F. J. 2019. Proteomic study of the membrane components of signalling cascades of Botrytis cinerea controlled by phosphorylation. *Scientific Reports,* 9**,** 9860.

FAN, G., YANG, Y., LI, T., LU, W., DU, Y., QIANG, X., WEN, Q. & SHAN, W. 2018. A Phytophthora capsici RXLR Effector Targets and Inhibits a Plant PPIase to Suppress Endoplasmic Reticulum-Mediated Immunity. *Mol Plant,* 11**,** 1067-1083.

FANG, Y., CUI, L., GU, B., ARREDONDO, F. & TYLER, B. M. 2017. Efficient Genome Editing in the Oomycete Phytophthora sojae Using CRISPR/Cas9. *Curr Protoc Microbiol,* 44**,** 21a.1.1-21a.1.26.

FANG, Y., GUPTA, V., KARRA, R., HOLDWAY, J. E., KIKUCHI, K. & POSS, K. D. 2013. Translational profiling of cardiomyocytes identifies an early Jak1/Stat3 injury response required for zebrafish heart regeneration. *Proc Natl Acad Sci U S A,* 110**,** 13416-21.

FANG, Y. & TYLER, B. M. 2016. Efficient disruption and replacement of an effector gene in the oomycete Phytophthora sojae using CRISPR/Cas9. 17**,** 127-39.

FARHANA, M. D. S. N. U. O. P. M., SARAWAK (MALAYSIA). DEPT. OF CROP SCIENCE), BIVI, M. R. U. O. P. M., SARAWAK (MALAYSIA). DEPT. OF CROP SCIENCE), KHAIRULMAZMI, A. U. O. P. M., SARAWAK (MALAYSIA). TROPICAL FOREST ECOSYSTEM SCIENCE), WONG, S. K. U. O. P. M., SARAWAK (MALAYSIA). DEPT. OF ANIMAL SCIENCE AND FISHERY) & SARIAH, M. U. O. P. M., SERDANG (MALAYSIA). DEPT. OF PLANT PROTECTION) 2013. Morphological and molecular characterization of Phytophthora capsici, the causal agent of foot rot disease of black pepper in Sarawak, Malaysia. v. 15.

GIANNAKOPOULOU, A., SCHORNACK, S., BOZKURT, T. O., HAART, D., RO, D.-K., FARALDOS, J. A., KAMOUN, S. & O'MAILLE, P. E. 2014. Variation in Capsidiol Sensitivity between Phytophthora infestans and Phytophthora capsici Is Consistent with Their Host Range. *PLOS ONE,* 9**,** e107462.

GIBSON, A. K. 2019. Asexual parasites and their extraordinary host ranges. *Integrative and Comparative Biology,* 59**,** 1463-1484.

GILBERT, G. S., MAGAREY, R., SUITER, K. & WEBB, C. O. 2012. Evolutionary tools for phytosanitary risk analysis: phylogenetic signal as a predictor of host range of plant pests and pathogens. *Evol Appl,* 5**,** 869-78.

GILBERT, G. S. & WEBB, C. O. 2007. Phylogenetic signal in plant pathogen-host range. *Proc Natl Acad Sci U S A,* 104**,** 4979-83.

GILBERT, K. J. & WHITLOCK, M. C. 2017. The genetics of adaptation to discrete heterogeneous environments: frequent mutation or large-effect alleles can allow range expansion. *Journal of Evolutionary Biology,* 30**,** 591-602.

GOSS, E. M., CARDENAS, M. E., MYERS, K., FORBES, G. A., FRY, W. E., RESTREPO, S. & GRÜNWALD, N. J. 2011. The plant pathogen Phytophthora andina emerged via hybridization of an unknown Phytophthora species and the Irish potato famine pathogen, P. infestans. *PLoS One,* 6**,** e24543.

GOSS, E. M., PRESS, C. M. & GRUNWALD, N. J. 2013. Evolution of RXLR-class effectors in the oomycete plant pathogen Phytophthora ramorum. *PLoS One,* 8**,** e79347.

GRACIDA, X. & CALARCO, J. A. 2017. Cell type-specific transcriptome profiling in C. elegans using the Translating Ribosome Affinity Purification technique. *Methods,* 126**,** 130-137.

GRENVILLE-BRIGGS, L. J., AVROVA, A. O., BRUCE, C. R., WILLIAMS, A., WHISSON, S. C., BIRCH, P. R. & VAN WEST, P. 2005. Elevated amino acid biosynthesis in Phytophthora infestans during appressorium formation and potato infection. *Fungal Genet Biol,* 42**,** 244-56.

GRUNWALD, N. J., GOSS, E. M. & PRESS, C. M. 2008. Phytophthora ramorum: a pathogen with a remarkably wide host range causing sudden oak death on oaks and ramorum blight on woody ornamentals. *Mol Plant Pathol,* 9**,** 729-40.

GUMTOW, R., WU, D., UCHIDA, J. & TIAN, M. 2018. A Phytophthora palmivora Extracellular Cystatin-Like Protease Inhibitor Targets Papain to Contribute to Virulence on Papaya. *Mol Plant Microbe Interact,* 31**,** 363-373.

GUO, M. & SCHIMMEL, P. 2013. Essential nontranslational functions of tRNA synthetases. *Nat Chem Biol,* 9**,** 145-53.

GUO, W., TZIOUTZIOU, N. A., STEPHEN, G., MILNE, I., CALIXTO, C. P., WAUGH, R., BROWN, J. W. S. & ZHANG, R. 2020. 3D RNA-seq: a powerful and flexible tool for rapid and accurate differential expression and alternative splicing analysis of RNA-seq data for biologists. *RNA Biol.***,** 1-14.

HAAS, B. J., KAMOUN, S., ZODY, M. C., JIANG, R. H., HANDSAKER, R. E., CANO, L. M., GRABHERR, M., KODIRA, C. D., RAFFAELE, S., TORTO-ALALIBO, T., BOZKURT, T. O., AH-FONG, A. M., ALVARADO, L., ANDERSON, V. L., ARMSTRONG, M. R., AVROVA, A., BAXTER, L., BEYNON, J., BOEVINK, P. C., BOLLMANN, S. R., BOS, J. I., BULONE, V., CAI, G., CAKIR, C., CARRINGTON, J. C., CHAWNER, M., CONTI, L., COSTANZO, S., EWAN, R., FAHLGREN, N., FISCHBACH, M. A., FUGELSTAD, J., GILROY, E. M., GNERRE, S., GREEN, P. J., GRENVILLE-BRIGGS, L. J., GRIFFITH, J., GRUNWALD, N. J., HORN, K., HORNER, N. R., HU, C. H., HUITEMA, E., JEONG, D. H., JONES, A. M., JONES, J. D., JONES, R. W., KARLSSON, E. K., KUNJETI, S. G., LAMOUR, K., LIU, Z., MA, L., MACLEAN, D., CHIBUCOS, M. C., MCDONALD, H., MCWALTERS, J., MEIJER, H. J., MORGAN, W., MORRIS, P. F., MUNRO, C. A., O'NEILL, K., OSPINA-GIRALDO, M., PINZON, A., PRITCHARD, L., RAMSAHOYE, B., REN, Q., RESTREPO, S., ROY, S., SADANANDOM, A., SAVIDOR, A., SCHORNACK, S., SCHWARTZ, D. C., SCHUMANN, U. D., SCHWESSINGER, B., SEYER, L., SHARPE, T., SILVAR, C., SONG, J., STUDHOLME, D. J., SYKES, S., THINES, M., VAN DE VONDERVOORT, P. J., PHUNTUMART, V., WAWRA, S., WEIDE, R., WIN, J., YOUNG, C., ZHOU, S., FRY, W., MEYERS, B. C., VAN WEST, P., RISTAINO, J., GOVERS, F., BIRCH, P. R., WHISSON, S. C., JUDELSON, H. S. & NUSBAUM, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. *Nature,* 461**,** 393-8.

HACKER, J., BLUM-OEHLER, G., MÜHLDORFER, I. & TSCHÄPE, H. 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol Microbiol,* 23**,** 1089-97.

HACQUARD, S., KRACHER, B., MAEKAWA, T., VERNALDI, S., SCHULZE-LEFERT, P. & VER LOREN VAN THEMAAT, E. 2013. Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. *Proceedings of the National Academy of Sciences,* 110**,** E2219.

HARRIS, L. J., BALCERZAK, M., JOHNSTON, A., SCHNEIDERMAN, D. & OUELLET, T. 2016. Host-preferential Fusarium graminearum gene expression during infection of wheat, barley, and maize. *Fungal Biol,* 120**,** 111-23.

HE, J., BALDINI, R. L., DÉZIEL, E., SAUCIER, M., ZHANG, Q., LIBERATI, N. T., LEE, D., URBACH, J., GOODMAN, H. M. & RAHME, L. G. 2004. The broad host range pathogen Pseudomonas

aeruginosa strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc Natl Acad Sci U S A,* 101**,** 2530-5.

HEIMAN, M., SCHAEFER, A., GONG, S., PETERSON, J. D., DAY, M., RAMSEY, K. E., SUAREZ-FARINAS, M., SCHWARZ, C., STEPHAN, D. A., SURMEIER, D. J., GREENGARD, P. & HEINTZ, N. 2008. A translational profiling approach for the molecular characterization of CNS cell types. *Cell,* 135**,** 738-48.

HOSSEINI, S., RESJÖ, S., LIU, Y., DURLING, M., HEYMAN, F., LEVANDER, F., LIU, Y., ELFSTRAND, M., FUNCK JENSEN, D., ANDREASSON, E. & KARLSSON, M. 2015. Comparative proteomic analysis of hyphae and germinating cysts of Phytophthora pisi and Phytophthora sojae. *J Proteomics,* 117**,** 24-40.

HSU, P. D., LANDER, E. S. & ZHANG, F. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell,* 157**,** 1262-1278.

HU, J., SHRESTHA, S., ZHOU, Y., MUDGE, J., LIU, X. & LAMOUR, K. 2020. Dynamic Extreme Aneuploidy (DEA) in the vegetable pathogen Phytophthora capsici and the potential for rapid asexual evolution. *PloS one,* 15**,** e0227250.

HU, X., XIAO, G., ZHENG, P., SHANG, Y., SU, Y., ZHANG, X., LIU, X., ZHAN, S., ST. LEGER, R. J. & WANG, C. 2014. Trajectory and genomic determinants of fungal-pathogen speciation and host adaptation. *Proceedings of the National Academy of Sciences,* 111**,** 16796.

HUANG, Y., NG, F. S. & JACKSON, F. R. 2015. Comparison of larval and adult Drosophila astrocytes reveals stage-specific gene expression profiles. *G3 (Bethesda),* 5**,** 551-8.

INOUE, Y., VY, T. T. P., YOSHIDA, K., ASANO, H., MITSUOKA, C., ASUKE, S., ANH, V. L., CUMAGUN, C. J. R., CHUMA, I., TERAUCHI, R., KATO, K., MITCHELL, T., VALENT, B., FARMAN, M. & TOSA, Y. 2017. Evolution of the wheat blast fungus through functional losses in a host specificity determinant. *Science,* 357**,** 80.

JIAO, Y. & MEYEROWITZ, E. M. 2010. Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. *Mol Syst Biol,* 6**,** 419.

JONES, J. D. & DANGL, J. L. 2006. The plant immune system. *Nature,* 444**,** 323-9.

JUDELSON, H. S. 1997. The genetics and biology of Phytophthora infestans: modern approaches to a historical challenge. *Fungal Genet Biol,* 22**,** 65-76.

JUDELSON, H. S. & SENTHIL, G. 2006. Investigating the role of ABC transporters in multifungicide insensitivity in Phytophthora infestans. *Mol Plant Pathol,* 7**,** 17-29.

JUPE, J., STAM, R., HOWDEN, A. J., MORRIS, J. A., ZHANG, R., HEDLEY, P. E. & HUITEMA, E. 2013. Phytophthora capsici-tomato interaction features dramatic shifts in gene expression associated with a hemi-biotrophic lifestyle. *Genome Biol,* 14**,** R63.

JWA, N. S. & HWANG, B. K. 2017. Convergent Evolution of Pathogen Effectors toward Reactive Oxygen Species Signaling Networks in Plants. *Front Plant Sci,* 8**,** 1687.

KAMOUN, S., FURZER, O., JONES, J. D., JUDELSON, H. S., ALI, G. S., DALIO, R. J., ROY, S. G., SCHENA, L., ZAMBOUNIS, A., PANABIERES, F., CAHILL, D., RUOCCO, M., FIGUEIREDO, A., CHEN, X. R., HULVEY, J., STAM, R., LAMOUR, K., GIJZEN, M., TYLER, B. M., GRUNWALD, N. J., MUKHTAR, M. S., TOME, D. F., TOR, M., VAN DEN ACKERVEKEN, G., MCDOWELL, J., DAAYF, F., FRY, W. E., LINDQVIST-KREUZE, H., MEIJER, H. J., PETRE, B., RISTAINO, J., YOSHIDA, K., BIRCH, P. R. & GOVERS, F. 2015. The Top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol,* 16**,** 413-34.

KASUGA, T. & GIJZEN, M. 2013. Epigenetics and the evolution of virulence. *Trends in Microbiology,* 21**,** 575-582.

KELLNER, R., BHATTACHARYYA, A., POPPE, S., HSU, T. Y., BREM, R. B. & STUKENBROCK, E. H. 2014. Expression profiling of the wheat pathogen Zymoseptoria tritici reveals genomic patterns of transcription and host-specific regulatory programs. *Genome Biol Evol,* 6**,** 1353-65.

KIRCHNER, J. W. & ROY, B. A. 2000. Evolutionary implications of host–pathogen specificity: the fitness consequences of host life history traits. *Evolutionary Ecology,* 14**,** 665-692.

KOU, Y. & NAQVI, N. I. 2016. Surface sensing and signaling networks in plant pathogenic fungi. *Semin Cell Dev Biol,* 57**,** 84-92.

LAMOUR, K. H., MUDGE, J., GOBENA, D., HURTADO-GONZALES, O. P., SCHMUTZ, J., KUO, A., MILLER, N. A., RICE, B. J., RAFFAELE, S., CANO, L. M., BHARTI, A. K., DONAHOO, R. S., FINLEY, S., HUITEMA, E., HULVEY, J., PLATT, D., SALAMOV, A., SAVIDOR, A., SHARMA, R., STAM, R., STOREY, D., THINES, M., WIN, J., HAAS, B. J., DINWIDDIE, D. L., JENKINS, J., KNIGHT, J. R., AFFOURTIT, J. P., HAN, C. S., CHERTKOV, O., LINDQUIST, E. A., DETTER, C., GRIGORIEV, I. V., KAMOUN, S. & KINGSMORE, S. F. 2012a. Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen Phytophthora capsici. *Mol Plant Microbe Interact,* 25**,** 1350-60.

LAMOUR, K. H., STAM, R., JUPE, J. & HUITEMA, E. 2012b. The oomycete broad-host-range pathogen Phytophthora capsici. *Mol Plant Pathol,* 13**,** 329-37.

LARSEN, M. K., JØRGENSEN, M. M., BENNIKE, T. B. & STENSBALLE, A. 2016. Time-course investigation of Phytophthora infestans infection of potato leaf from three cultivars by quantitative proteomics. *Data Brief,* 6**,** 238-48.

LEONIAN, L. H. 1922. Stem and fruit blight of Peppers caused by Phytophthora capsici sp. nov. *Phytopathology* 12**,** 401-408.

LEWIS, D. H. 1973. CONCEPTS IN FUNGAL NUTRITION AND THE ORIGIN OF BIOTROPHY. *Biological Reviews,* 48**,** 261-277.

LI, J., CORNELISSEN, B. & REP, M. 2020. Host-specificity factors in plant pathogenic fungi. *Fungal Genetics and Biology,* 144**,** 103447.

LIANG, X. & ROLLINS, J. A. 2018. Mechanisms of Broad Host Range Necrotrophic Pathogenesis in Sclerotinia sclerotiorum. 108**,** 1128-1140.

LIN, X., SONG, T., FAIRHEAD, S., WITEK, K., JOUET, A., JUPE, F., WITEK, A. I., KARKI, H. S., VLEESHOUWERS, V. G. A. A., HEIN, I. & JONES, J. D. G. 2020. Identification of Avramr1 from Phytophthora infestans using long read and cDNA pathogen-enrichment sequencing (PenSeq). *Molecular Plant Pathology,* 21**,** 1502-1512.

LINDEBERG, M., MYERS, C. R., COLLMER, A. & SCHNEIDER, D. J. 2008. Roadmap to new virulence determinants in Pseudomonas syringae: insights from comparative genomics and genome organization. *Mol Plant Microbe Interact,* 21**,** 685-700.

LIU, W., ZHOU, X., LI, G., LI, L., KONG, L., WANG, C., ZHANG, H. & XU, J. R. 2011. Multiple plant surface signals are sensed by different mechanisms in the rice blast fungus for appressorium formation. *PLoS Pathog,* 7**,** e1001261.

LOVELL, H. C., MANSFIELD, J. W., GODFREY, S. A., JACKSON, R. W., HANCOCK, J. T. & ARNOLD, D. L. 2009. Bacterial evolution by genomic island transfer occurs via DNA transformation in planta. *Curr Biol,* 19**,** 1586-90.

LUO, Q., WANG, F. X., ZHONG, N. Q., WANG, H. Y. & XIA, G. X. 2014. The role of autophagy during development of the oomycete pathogen Phytophthora infestans. *J Genet Genomics,* 41**,** 225-8.

LYONS, L. C., CHATTERJEE, S., VANROBAEYS, Y., GAINE, M. E. & ABEL, T. 2020. Translational changes induced by acute sleep deprivation uncovered by TRAP-Seq. *Molecular Brain,* 13**,** 165.

MA, H., SHEN, D., WU, Y., XU, H. & DOU, D. 2018. RNA-seq for comparative transcript profiling of Phytophthora capsici during its interaction with Arabidopsis thaliana. *Physiological and Molecular Plant Pathology,* 102**,** 193-199.

MA, L. J., VAN DER DOES, H. C., BORKOVICH, K. A., COLEMAN, J. J., DABOUSSI, M. J., DI PIETRO, A., DUFRESNE, M., FREITAG, M., GRABHERR, M., HENRISSAT, B., HOUTERMAN, P. M., KANG, S., SHIM, W. B., WOLOSHUK, C., XIE, X., XU, J. R., ANTONIW, J., BAKER, S. E., BLUHM, B. H., BREAKSPEAR, A., BROWN, D. W., BUTCHKO, R. A., CHAPMAN, S., COULSON, R., COUTINHO, P. M., DANCHIN, E. G., DIENER, A., GALE, L. R., GARDINER, D. M., GOFF, S., HAMMOND-KOSACK, K. E., HILBURN, K., HUA-VAN, A., JONKERS, W., KAZAN, K., KODIRA, C. D., KOEHRSEN, M., KUMAR, L., LEE, Y. H., LI, L., MANNERS, J. M., MIRANDA-SAAVEDRA, D.,

MUKHERJEE, M., PARK, G., PARK, J., PARK, S. Y., PROCTOR, R. H., REGEV, A., RUIZ-ROLDAN, M. C., SAIN, D., SAKTHIKUMAR, S., SYKES, S., SCHWARTZ, D. C., TURGEON, B. G., WAPINSKI, I., YODER, O., YOUNG, S., ZENG, Q., ZHOU, S., GALAGAN, J., CUOMO, C. A., KISTLER, H. C. & REP, M. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature,* 464**,** 367-73.

MA, Z., ZHU, L., SONG, T., WANG, Y., ZHANG, Q., XIA, Y., QIU, M., LIN, Y., LI, H., KONG, L., FANG, Y., YE, W., WANG, Y., DONG, S., ZHENG, X., TYLER, B. M. & WANG, Y. 2017. A paralogous decoy protects Phytophthora sojae apoplastic effector PsXEG1 from a host inhibitor. *Science,* 355**,** 710.

MBENGUE, M., NAVAUD, O., PEYRAUD, R., BARASCUD, M., BADET, T., VINCENT, R., BARBACCI, A. & RAFFAELE, S. 2016. Emerging Trends in Molecular Interactions between Plants and the Broad Host Range Fungal Pathogens Botrytis cinerea and Sclerotinia sclerotiorum. *Front Plant Sci,* 7**,** 422.

MEILE, L., PETER, J., PUCCETTI, G., ALASSIMONE, J., MCDONALD, B. A. & SÁNCHEZ-VALLET, A. 2020. Chromatin Dynamics Contribute to the Spatiotemporal Expression Pattern of Virulence Genes in a Fungal Plant Pathogen. *mBio,* 11**,** e02343-20.

MENARDO, F., PRAZ, C. R., WYDER, S., BEN-DAVID, R., BOURRAS, S., MATSUMAE, H., MCNALLY, K. E., PARLANGE, F., RIBA, A., ROFFLER, S., SCHAEFER, L. K., SHIMIZU, K. K., VALENTI, L., ZBINDEN, H., WICKER, T. & KELLER, B. 2016. Hybridization of powdery mildew strains gives rise to pathogens on novel agricultural crop species. *Nat Genet,* 48**,** 201-5.

MIAO, J., CHI, Y., LIN, D., TYLER, B. M. & LIU, X. 2018. Mutations in ORP1 Conferring Oxathiapiprolin Resistance Confirmed by Genome Editing using CRISPR/Cas9 in Phytophthora capsici and P. sojae. *Phytopathology,* 108**,** 1412-1419.

MORRIS, C. E. & MOURY, B. 2019. Revisiting the Concept of Host Range of Plant Pathogens. *Annu Rev Phytopathol,* 57**,** 63-90.

MORRIS, P. F. & PHUNTUMART, V. 2009. Inventory and comparative evolution of the ABC superfamily in the genomes of Phytophthora ramorum and Phytophthora sojae. *J Mol Evol,* 68**,** 563-75.

NAKAO, A., YOSHIHAMA, M. & KENMOCHI, N. 2004. RPG: the Ribosomal Protein Gene database. *Nucleic Acids Res,* 32**,** D168-70.

NÄPFLIN, K., O'CONNOR, E. A., BECKS, L., BENSCH, S., ELLIS, V. A., HAFER-HAHMANN, N., HARDING, K. C., LINDÉN, S. K., OLSEN, M. T., ROVED, J., SACKTON, T. B., SHULTZ, A. J., VENKATAKRISHNAN, V., VIDEVALL, E., WESTERDAHL, H., WINTERNITZ, J. C. & EDWARDS, S. V. 2019. Genomics of host-pathogen interactions: challenges and opportunities across ecological and spatiotemporal scales. *PeerJ,* 7**,** e8013-e8013.

NASS 2020. Vegetables 2019 Summary. *In:* AGRICULTURE, N. A. S. S.-U. S. D. O. (ed.).

NEWMAN, T. E. & DERBYSHIRE, M. C. 2020. The Evolutionary and Molecular Features of Broad Host-Range Necrotrophy in Plant Pathogenic Fungi. *Frontiers in plant science,* 11**,** 591733-591733.

OLIVER, R. P. & IPCHO, S. V. 2004. Arabidopsis pathology breathes new life into the necrotrophs-vs.-biotrophs classification of fungal pathogens. *Mol Plant Pathol,* 5**,** 347-52.

OLIVER, R. P. & SOLOMON, P. S. 2010a. New developments in pathogenicity and virulence of necrotrophs. *Curr Opin Plant Biol,* 13**,** 415-9.

OLIVER, R. P. & SOLOMON, P. S. 2010b. New developments in pathogenicity and virulence of necrotrophs. *Current Opinion in Plant Biology,* 13**,** 415-419.

PANG, Z., CHEN, L., MU, W., LIU, L. & LIU, X. 2016. Insights into the adaptive response of the plant-pathogenic oomycete Phytophthora capsici to the fungicide flumorph. *Scientific Reports,* 6**,** 24103.

PANG, Z., SRIVASTAVA, V., LIU, X. & BULONE, V. 2017. Quantitative proteomics links metabolic pathways to specific developmental stages of the plant-pathogenic oomycete Phytophthora capsici. *Mol Plant Pathol,* 18**,** 378-390.

PANSTRUGA, R. & MOSCOU, M. J. 2020. What is the Molecular Basis of Nonhost Resistance? *Molecular Plant-Microbe Interactions®,* 33**,** 1253-1264.

PARK, C. J. & SEO, Y. S. 2015. Heat Shock Proteins: A Review of the Molecular Chaperones for Plant Immunity. *Plant Pathol J,* 31**,** 323-33.

PARKHILL, J., DOUGAN, G., JAMES, K. D., THOMSON, N. R., PICKARD, D., WAIN, J., CHURCHER, C., MUNGALL, K. L., BENTLEY, S. D., HOLDEN, M. T. G., SEBAIHIA, M., BAKER, S., BASHAM, D., BROOKS, K., CHILLINGWORTH, T., CONNERTON, P., CRONIN, A., DAVIS, P., DAVIES, R. M., DOWD, L., WHITE, N., FARRAR, J., FELTWELL, T., HAMLIN, N., HAQUE, A., HIEN, T. T., HOLROYD, S., JAGELS, K., KROGH, A., LARSEN, T. S., LEATHER, S., MOULE, S., Ó'GAORA, P., PARRY, C., QUAIL, M., RUTHERFORD, K., SIMMONDS, M., SKELTON, J., STEVENS, K., WHITEHEAD, S. & BARRELL, B. G. 2001a. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. *Nature,* 413**,** 848-852.

PARKHILL, J., WREN, B. W., THOMSON, N. R., TITBALL, R. W., HOLDEN, M. T., PRENTICE, M. B., SEBAIHIA, M., JAMES, K. D., CHURCHER, C., MUNGALL, K. L., BAKER, S., BASHAM, D., BENTLEY, S. D., BROOKS, K., CERDEÑO-TÁRRAGA, A. M., CHILLINGWORTH, T., CRONIN, A., DAVIES, R. M., DAVIS, P., DOUGAN, G., FELTWELL, T., HAMLIN, N., HOLROYD, S., JAGELS, K., KARLYSHEV, A. V., LEATHER, S., MOULE, S., OYSTON, P. C., QUAIL, M., RUTHERFORD, K., SIMMONDS, M., SKELTON, J., STEVENS, K., WHITEHEAD, S. & BARRELL, B. G. 2001b. Genome sequence of Yersinia pestis, the causative agent of plague. *Nature,* 413**,** 523-7.

PATRO, R., DUGGAL, G., LOVE, M. I., IRIZARRY, R. A. & KINGSFORD, C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods,* 14**,** 417-419.

PENG, D., KURUP, S. P., YAO, P. Y., MINNING, T. A. & TARLETON, R. L. 2014. CRISPR-Cas9-mediated single-gene and gene family disruption in Trypanosoma cruzi. *mBio,* 6**,** e02097-14.

PERFECT, S. E. & GREEN, J. R. 2001. Infection structures of biotrophic and hemibiotrophic fungal plant pathogens. *Mol Plant Pathol,* 2**,** 101-8.

PETTONGKHAO, S., NAVET, N., SCHORNACK, S., TIAN, M. & CHURNGCHOW, N. 2020. A secreted protein of 15 kDa plays an important role in Phytophthora palmivora development and pathogenicity. *Scientific Reports,* 10**,** 2319.

PHAM, J., STAM, R., HEREDIA, V. M., CSUKAI, M. & HUITEMA, E. 2018. An NMRA-Like Protein Regulates Gene Expression in Phytophthora capsici to Drive the Infection Cycle on Tomato. *Mol Plant Microbe Interact,* 31**,** 665-677.

PIASECKA, A., JEDRZEJCZAK-REY, N. & BEDNAREK, P. 2015. Secondary metabolites in plant innate immunity: conserved function of divergent chemicals. *New Phytol,* 206**,** 948-64.

PUIG, A. S., ALI, S., STREM, M., SICHER, R., GUTIERREZ, O. A. & BAILEY, B. A. 2018. The differential influence of temperature on Phytophthora megakarya and Phytophthora palmivora pod lesion expansion, mycelia growth, gene expression, and metabolite profiles. *Physiological and Molecular Plant Pathology,* 102**,** 95-112.

QUANDT, C. A., DI, Y., ELSER, J., JAISWAL, P. & SPATAFORA, J. W. 2016. Differential Expression of Genes Involved in Host Recognition, Attachment, and Degradation in the Mycoparasite Tolypocladium ophioglossoides. *G3 (Bethesda),* 6**,** 731-41.

RAFFAELE, S., FARRER, R. A., CANO, L. M., STUDHOLME, D. J., MACLEAN, D., THINES, M., JIANG, R. H., ZODY, M. C., KUNJETI, S. G., DONOFRIO, N. M., MEYERS, B. C., NUSBAUM, C. & KAMOUN, S. 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science,* 330**,** 1540-3.

RESJÖ, S., BRUS, M., ALI, A., MEIJER, H. J. G., SANDIN, M., GOVERS, F., LEVANDER, F., GRENVILLE-BRIGGS, L. & ANDREASSON, E. 2017. Proteomic Analysis of Phytophthora infestans Reveals the Importance of Cell Wall Proteins in Pathogenicity. *Mol Cell Proteomics,* 16**,** 1958-1971.

REYNOSO, M. A., JUNTAWONG, P., LANCIA, M., BLANCO, F. A., BAILEY-SERRES, J. & ZANETTI, M. E. 2015. Translating Ribosome Affinity Purification (TRAP) followed by RNA sequencing technology (TRAP-SEQ) for quantitative assessment of plant translatomes. *Methods Mol Biol,* 1284**,** 185-207.

ROBERTS, P. D., URS, R. R., FRENCH-MONAR, R. D., HOFFINE, M. S., SEIJO, T. E. & MCGOVERN, R. J. 2005. Survival and recovery of Phytophthora capsici and oomycetes in tailwater and soil from vegetable fields in Florida. *Annals of Applied Biology,* 146**,** 351-359.

RODENBURG, S. Y. A., SEIDL, M. F., JUDELSON, H. S., VU, A. L., GOVERS, F. & DE RIDDER, D. 2019. Metabolic Model of the Phytophthora infestans-Tomato Interaction Reveals Metabolic Switches during Host Colonization. *mBio,* 10**,** e00454-19.

ROSS, L., HARDY, N. B., OKUSU, A. & NORMARK, B. B. 2013. Large population size predicts the distribution of asexuality in scale insects. *Evolution,* 67**,** 196-206.

SAVARY, S., WILLOCQUET, L., PETHYBRIDGE, S. J., NELSON, A., MCROBERTS, N. & ESKER, P. 2019. The global burden of pathogens and pests on major food crops. *Nat Ecol Evol,* 3**,** 430-439.

SCHLAEPPI, K. & MAUCH, F. 2010. Indolic secondary metabolites protect Arabidopsis from the oomycete pathogen Phytophthora brassicae. *Plant Signal Behav,* 5**,** 1099-101.

SCHNITTER, R. & BERRY, P. 2019. The Climate Change, Food Security and Human Health Nexus in Canada: A Framework to Protect Population Health. *Int J Environ Res Public Health,* 16.

SCHOINA, C., BOUWMEESTER, K. & GOVERS, F. 2017. Infection of a tomato cell culture by Phytophthora infestans; a versatile tool to study Phytophthora-host interactions. *Plant Methods,* 13**,** 88.

SCHULZE-LEFERT, P. & PANSTRUGA, R. 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. *Trends Plant Sci,* 16**,** 117-25.

SHARMA, R., MISHRA, B., RUNGE, F. & THINES, M. 2014. Gene loss rather than gene gain is associated with a host jump from monocots to dicots in the Smut Fungus Melanopsichium pennsylvanicum. *Genome Biol Evol,* 6**,** 2034-49.

SILVA-BELTRÁN, N. P., RUIZ-CRUZ, S., CIRA-CHÁVEZ, L. A., ESTRADA-ALVARADO, M. I., ORNELAS-PAZ, J. D. J., LÓPEZ-MATA, M. A., DEL-TORO-SÁNCHEZ, C. L., AYALA-ZAVALA, J. F. & MÁRQUEZ-RÍOS, E. 2015. Total Phenolic, Flavonoid, Tomatine, and Tomatidine Contents and Antioxidant and Antimicrobial Activities of Extracts of Tomato Plant. *International Journal of Analytical Chemistry,* 2015**,** 284071.

SLANE, D., KONG, J., BERENDZEN, K. W., KILIAN, J., HENSCHEN, A., KOLB, M., SCHMID, M., HARTER, K., MAYER, U., DE SMET, I., BAYER, M. & JURGENS, G. 2014. Cell type-specific transcriptome analysis in the early Arabidopsis thaliana embryo. *Development,* 141**,** 4831-40.

SRIVASTAVA, A., CHO, I. K. & CHO, Y. 2013. The Bdtf1 gene in Alternaria brassicicola is important in detoxifying brassinin and maintaining virulence on Brassica species. *Mol Plant Microbe Interact,* 26**,** 1429-40.

STAAL, J., KALIFF, M., BOHMAN, S. & DIXELIUS, C. 2006. Transgressive segregation reveals two Arabidopsis TIR-NB-LRR resistance genes effective against Leptosphaeria maculans, causal agent of blackleg disease. *Plant J,* 46**,** 218-30.

STAM, R., JUPE, J., HOWDEN, A. J., MORRIS, J. A., BOEVINK, P. C., HEDLEY, P. E. & HUITEMA, E. 2013. Identification and Characterisation CRN Effectors in Phytophthora capsici Shows Modularity and Functional Diversity. *PLoS One,* 8**,** e59517.

STAM, R., MANTELIN, S., MCLELLAN, H. & THILLIEZ, G. 2014. The role of effectors in nonhost resistance to filamentous plant pathogens. *Front Plant Sci,* 5**,** 582.

SUBIN, C. S., PRADEEP, M. A. & VIJAYAN, K. K. 2016. FKBP-type peptidyl-prolyl cis-trans isomerase from thermophilic microalga, Scenedesmus sp.: molecular characterisation and demonstration of acquired salinity and thermotolerance in E. coli by recombinant expression. *Journal of Applied Phycology,* 28**,** 3307-3315.

SUSI, H., BURDON, J. J., THRALL, P. H., NEMRI, A. & BARRETT, L. G. 2020. Genetic analysis reveals long-standing population differentiation and high diversity in the rust pathogen Melampsora lini. *PLOS Pathogens,* 16**,** e1008731.

TAKEUCHI, K., TAGUCHI, F., INAGAKI, Y., TOYODA, K., SHIRAISHI, T. & ICHINOSE, Y. 2003. Flagellin glycosylation island in Pseudomonas syringae pv. glycinea and its role in host specificity. *Journal of bacteriology,* 185**,** 6658-6665.

THOMAS, A., LEE, P.-J., DALTON, J. E., NOMIE, K. J., STOICA, L., COSTA-MATTIOLI, M., CHANG, P., NUZHDIN, S., ARBEITMAN, M. N. & DIERICK, H. A. 2012. A Versatile Method for Cell-Specific Profiling of Translated mRNAs in Drosophila. *PLOS ONE,* 7**,** e40276.

TIAN, D. & BABADOOST, M. 2004. Host Range of Phytophthora capsici from Pumpkin and Pathogenicity of Isolates. *Plant Dis,* 88**,** 485-489.

TRYON, R. C., PISAT, N., JOHNSON, S. L. & DOUGHERTY, J. D. 2013. Development of translating ribosome affinity purification for zebrafish. *Genesis,* 51**,** 187-92.

TZELEPIS, G., HODÉN, K. P., FOGELQVIST, J., ÅSMAN, A. K. M., VETUKURI, R. R. & DIXELIUS, C. 2020. Dominance of Mating Type A1 and Indication of Epigenetic Effects During Early Stages of Mating in Phytophthora infestans. *Front Microbiol,* 11**,** 252.

VAN DEN HOOGEN, J. & GOVERS, F. 2018. Attempts to implement CRISPR/Cas9 for genome editing in the oomycete Phytophthora infestans</em&gt. *bioRxiv***,** 274829.

VAN WEST, P., SHEPHERD, S. J., WALKER, C. A., LI, S., APPIAH, A. A., GRENVILLE-BRIGGS, L. J., GOVERS, F. & GOW, N. A. R. 2008. Internuclear gene silencing in Phytophthora infestans is established through chromatin remodelling. *Microbiology (Reading),* 154**,** 1482-1490.

VEGA-ARREGUIN, J. C., SHIMADA-BELTRAN, H., SEVILLANO-SERRANO, J. & MOFFETT, P. 2017. Non-host Plant Resistance against Phytophthora capsici Is Mediated in Part by Members of the I2 R Gene Family in Nicotiana spp. *Front Plant Sci,* 8**,** 205.

VIJN, I. & GOVERS, F. 2003. Agrobacterium tumefaciens mediated transformation of the oomycete plant pathogen Phytophthora infestans. *Mol Plant Pathol,* 4**,** 459-67.

WANG, Q., HAN, C., FERREIRA, A. O., YU, X., YE, W., TRIPATHY, S., KALE, S. D., GU, B., SHENG, Y., SUI, Y., WANG, X., ZHANG, Z., CHENG, B., DONG, S., SHAN, W., ZHENG, X., DOU, D., TYLER, B. M. & WANG, Y. 2011. Transcriptional Programming and Functional Interactions within the Phytophthora sojae RXLR Effector Repertoire. *The Plant Cell,* 23**,** 2064.

WANG, W., XUE, Z., MIAO, J., CAI, M., ZHANG, C., LI, T., ZHANG, B., TYLER, B. M. & LIU, X. 2019. PcMuORP1, an Oxathiapiprolin-Resistance Gene, Functions as a Novel Selection Marker for Phytophthora Transformation and CRISPR/Cas9 Mediated Genome Editing. *Frontiers in microbiology,* 10**,** 2402-2402.

WANG, Z., TYLER, B. M. & LIU, X. 2018. Protocol of Phytophthora capsici Transformation Using the CRISPR-Cas9 System. *Methods Mol Biol,* 1848**,** 265-274.

WARD, B., VAN OOSTERHOUT, C., THILLIEZ, G. J. A., ARMSTRONG, M. R., LIM, T. Y., BAKER, K., JOUET, A., JONES, J. D. G., HUITEMA, E., HEIN, I. & BIRCH, P. R. J. 2019. Pathogen enrichment sequencing (PenSeq) enables population genomic studies in oomycetes. *New Phytol,* 221**,** 1634-1648.

WARD, E. W. B. & STOESSL, A. 1972. Postinfectional Inhibitors from Plants. III. Detoxification of Capsidiol, an Antifungal Compound from Peppers. *Phytopathology,* 62**,** 1186-1187.

WATSON, F. L., MILLS, E. A., WANG, X., GUO, C., CHEN, D. F. & MARSH-ARMSTRONG, N. 2012. Cell type–specific translational profiling in the Xenopus laevis retina. *Developmental Dynamics,* 241**,** 1960-1972.

WHISSON, S., VETUKURI, R., AVROVA, A. & DIXELIUS, C. 2012. Can silencing of transposons contribute to variation in effector gene expression in Phytophthora infestans? *Mobile genetic elements,* 2**,** 110-114.

WHISSON, S. C., AVROVA, A. O., P, V. A. N. W. & JONES, J. T. 2005. A method for double-stranded RNA-mediated transient gene silencing in Phytophthora infestans. *Mol Plant Pathol,* 6**,** 153-63.

WHISSON, S. C., BOEVINK, P. C., MOLELEKI, L., AVROVA, A. O., MORALES, J. G., GILROY, E. M., ARMSTRONG, M. R., GROUFFAUD, S., VAN WEST, P., CHAPMAN, S., HEIN, I., TOTH, I. K., PRITCHARD, L. & BIRCH, P. R. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature,* 450**,** 115-8.

WU, D., NAVET, N., LIU, Y., UCHIDA, J. & TIAN, M. 2016. Establishment of a simple and efficient Agrobacterium-mediated transformation system for Phytophthora palmivora. *BMC Microbiology,* 16**,** 204.

XIAO, C., GAO, J., ZHANG, Y., WANG, Z., ZHANG, D., CHEN, Q., YE, X., XU, Y., YANG, G., YAN, L., CHENG, Q., CHEN, J. & SHEN, Y. 2019. Quantitative Proteomics of Potato Leaves Infected with Phytophthora infestans Provides Insights into Coordinated and Altered Protein Expression during Early and Late Disease Stages. *Int J Mol Sci,* 20.

YAN, S., LIU, H., MOHR, T. J., JENRETTE, J., CHIODINI, R., ZACCARDELLI, M., SETUBAL, J. C. & VINATZER, B. A. 2008. Role of recombination in the evolution of the model plant pathogen Pseudomonas syringae pv. tomato DC3000, a very atypical tomato strain. *Applied and environmental microbiology,* 74**,** 3171-3181.

YANG, M., DUAN, S., MEI, X., HUANG, H., CHEN, W., LIU, Y., GUO, C., YANG, T., WEI, W., LIU, X., HE, X., DONG, Y. & ZHU, S. 2018. The Phytophthora cactorum genome provides insights into the adaptation to host defense compounds and fungicides. *Scientific Reports,* 8**,** 6534.

YOON, B. C., JUNG, H., DWIVEDY, A., O'HARE, C. M., ZIVRAJ, K. H. & HOLT, C. E. 2012. Local translation of extranuclear lamin B promotes axon maintenance. *Cell,* 148**,** 752-64.

YUNUSA, A. K., DANDAGO, M. A., IBRAHIM, S. A. M., ABDULLAHI, N., RILWAN, A. & BARDE, A. 2018. Total Phenolic Content and Antioxidant Capacity of Different Parts of Cucumber (Cucumis sativus L.). *Acta Universitatis Cibiniensis. Series E: Food Technology,* 22**,** 13-20.

ZHANG, C., XIAO, B., JIANG, Y., ZHAO, Y., LI, Z., GAO, H., LING, Y., WEI, J., LI, S., LU, M., SU, X. Z., CUI, H. & YUAN, J. 2014. Efficient editing of malaria parasite genome using the CRISPR/Cas9 system. *mBio,* 5**,** e01414-14.

ZHEN, Y. & STENMARK, H. 2015. Cellular functions of Rab GTPases at a glance. *J Cell Sci,* 128**,** 3171-6.

ZHENG, X., MCLELLAN, H., FRAITURE, M., LIU, X., BOEVINK, P. C., GILROY, E. M., CHEN, Y., KANDEL, K., SESSA, G., BIRCH, P. R. & BRUNNER, F. 2014. Functionally redundant RXLR effectors from Phytophthora infestans act at different steps to suppress early flg22-triggered immunity. *PLoS Pathog,* 10**,** e1004057.

Appendix

Supplementary Table 1: Cross referencing DE genes in each cluster group with the comparison group that the are DE in.

The number of genes from each of the 10 clusters form the hierarchical cluster analysis for the all gene database and the number of those gene from comparison group and each combination of groups where genes occur in are shown here. The total number of gene in each cluster in shown down the right and the total number of genes in each comparison group and each combination of groups are shown across the top. The comparison groups are separated into the three extract vs germinating (Extract v GC: TE 2hpi, CE 2hpi and Extract 2hpi (Ex)) and the four comparison of the tomato extract and cucumber extract (TOM v CUC: 2 hours post inoculation (hpi), 4 hpi, 8hpi and All hpi), and a combination of these. For the genes in Extract vs GC those labelled as Down are up regulated in GC and those labelled as Up are up regulated in Extract. For the genes in TOM v CUC those labelled as Down are up regulated in CUC and those labelled as Up are up regulated in TOM.

		Extract vs GC												Tom vs Cuc							
		2 hpi TE		2 hpi TE + Ex		2 hpi Ex		2 hpi CE + Ex		2 hpi TE + CE + Ex		2 hpi CE		All hpi		2 + All hpi	2 hpi		8 hpi	2 + 4 + 8 + All hpi	2 + 4 + All hpi
		Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Down	Up	Up	Up	Up
Cluster		150	21	200	53	190	42	75	68	43	133	7	54	38	23	22	50	3	1	1	1
1	40		14		6				1		11		1		10			2		1	
2	58		1		1		5		12		37		1	2			1				
3	106			4		46		44		3		7			2			1			1
4	125	5		1			2		33		11		45	18		8	18				
5	216	5		70		83		25		33				3							
6	102				17		20		14		45		6								
7	218	45		97		61		5		7				5		3	4				
8	96	2	6		29		15		8		29		1		11				1		
9	85	58		18										9		6	12				
10	46	35		10				1						1		5	15				

Supplementary Table 2: Cross referencing DE of genes from every comparison group.

The comparison groups are separated into the three extract vs germinating (Extract v GC: TE 2hpi, CE 2hpi and Extract 2hpi (Ex)) and the four comparison of the tomato extract and cucumber extract (TE v CE: 2 hours post inoculation (hpi), 4 hpi, 8hpi and All hpi), and a combination of these. Differentially expressed (DE) gene, are either thus found DE in a single or multiple (combination) of comparisons. For the genes in Extract vs GC those labelled as Down are up regulated in GC and those labelled as Up are up regulated in Extract. For the genes in TE v CE those labelled as Down are up regulated in CE and those labelled as Up are up regulated in TE. Note that not all the genes from each comparison group of TE v CE appears in extract v GC and vis versa.

				Extract v GC											
				2 hpi TE		2 hpi TE + Ex		2 hpi Ex		2 hpi CE + Ex		2 hpi TE + CE + Ex		2 hpi CE	
				Down	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down	Up
				150	21	200	53	190	42	75	68	43	133	7	54
TE v CE	All hpi	Down	38	**1**		**4**						**2**	**1**		**3**
		Up	23		**3**		**2**						**4**		
	2 + All hpi	Down	22	**7**		**10**							**1**		
	2 hpi	Down	50	**22**		**11**					**2**		**3**		**2**
		Up	3	**2**										**1**	
	8 hpi	Up	1												
	2 + 4 + 8 + All hpi	Up	1		**1**										
	2 + 4 + All hpi	Up	1							**1**					

Supplementary Table 3: GO terms and KOG annotations for all DE proteins from Extract versus GC comparisons

Phyca 11 number	Gene Ontology	KOG
Induced in EX		
Phyca11_104444	GO:0005524 ATP binding	Molecular chaperones HSP70/HSC70, HSP70 superfamily
Phyca11_113258	-	26S proteasome regulatory complex, subunit RPN11
Phyca11_21374	GO:0003735 structural constituent of ribosome, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006412 translation	40S ribosomal protein S14
Phyca11_40390	GO:0003735 structural constituent of ribosome, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006412 translation	60s ribosomal protein L19
Phyca11_502611	-	eneurin-1 and related extracellular matrix proteins, contain EGF-like repeats
Phyca11_503054	-	NADH:ubiquinone oxidoreductase, NDUFA6/B14 subunit
Phyca11_503551	GO:0003824 catalytic activity, GO:0005524 ATP binding, GO:0016301 kinase activity, GO:0016310 phosphorylation, GO:0016772 transferase activity, transferring phosphorus-containing groups, GO:0050242 pyruvate, phosphate dikinase activity	-
Phyca11_503721	-	-
Phyca11_503953	GO:0003924 GTPase activity, GO:0005515 protein binding, GO:0005524 ATP binding, GO:0005525 GTP binding, GO:0005622 intracellular, GO:0006355 regulation of transcription, DNA-templated, GO:0006886 intracellular protein transport, GO:0006913 nucleocytoplasmic transport, GO:0007165 signal transduction, GO:0007264 small GTPase mediated signal transduction, GO:0008134 transcription factor binding, GO:0015031 protein transport	GTPase Rab11/YPT3, small G protein superfamily
Phyca11_503954	GO:0003676 nucleic acid binding	Phosphoprotein/predicted coiled-coil protein
Phyca11_504023	GO:0003676 nucleic acid binding, GO:0003735 structural constituent of ribosome, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006412 translation, GO:0006464 cellular protein modification process, GO:0008270 zinc ion ,	Ubiquitin/40S ribosomal protein S27a fusion
Phyca11_504650	GO:0003924 GTPase activity, GO:0005515 protein binding, GO:0005524 ATP binding, GO:0005525 GTP binding, GO:0005622 intracellular, GO:0006355 regulation of transcription, DNA-templated, GO:0006886 intracellular protein transport, GO:0006913 nucleocytoplasmic transport, GO:0007165 signal transduction, GO:0007264 small GTPase mediated signal transduction, GO:0008134 transcription factor binding, GO:0015031 protein transport	GTPase Rab2, small G protein superfamily
Phyca11_505507	GO:0005509 calcium ion binding, GO:0005544 calcium-dependent phospholipid binding	Annexin
Phyca11_509200	GO:0003824 catalytic activity, GO:0005488 binding, GO:0008152 metabolic process ,	Predicted NAD-dependent oxidoreductase
Phyca11_509792	GO:0003735 structural constituent of ribosome, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006414 translational elongation	60S acidic ribosomal protein P2
Phyca11_510076	GO:0006457 protein folding	FKBP-type peptidyl-prolyl cis-trans isomerase
Phyca11_510914	GO:0003735 structural constituent of ribosome, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006412 translation	40S ribosomal protein S13
Phyca11_511272	GO:0003924 GTPase activity, GO:0005198 structural molecule activity, GO:0005525 GTP binding, GO:0005874 microtubule, GO:0007017 microtubule-based process, GO:0007018 microtubule-based movement, GO:0043234 protein complex, GO:0051258 protein polymerization	Alpha tubulin
Phyca11_512003	GO:0015986 ATP synthesis coupled proton transport, GO:0016469 proton-transporting two-sector ATPase complex,	Mitochondrial F1F0-ATP synthase, subunit delta/ATP16

	GO:0046933 proton-transporting ATP synthase activity, rotational mechanism, GO:0046961 proton-transporting ATPase activity, rotational mechanism	
Phyca11_528971	GO:0003824 catalytic activity, GO:0005506 iron ion binding, GO:0006725 cellular aromatic compound metabolic process, GO:0008199 ferric iron binding, GO:0016702 oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen ,	Dioxygenase
Phyca11_530856	GO:0003924 GTPase activity, GO:0005198 structural molecule activity, GO:0005525 GTP binding, GO:0005874 microtubule, GO:0007017 microtubule-based process, GO:0007018 microtubule-based movement, GO:0043234 protein complex, GO:0051258 protein polymerization	Alpha tubulin
Phyca11_534612	GO:0003824 catalytic activity, GO:0005975 carbohydrate metabolic process, GO:0006013 mannose metabolic process, GO:0015923 mannosidase activity, GO:0030246 carbohydrate binding ,	Glycosyl hydrolase, family 38 - alpha-mannosidase
Phyca11_536157	GO:0004618 phosphoglycerate kinase activity, GO:0006096 glycolytic process	3-phosphoglycerate kinase
Phyca11_538207	GO:0000166 nucleotide binding, GO:0003676 nucleic acid binding, GO:0004386 helicase activity, GO:0005524 ATP binding, GO:0008026 ATP-dependent helicase activity, GO:0017111 nucleoside-triphosphatase activity	DEAH-box RNA helicase
Phyca11_538675	GO:0000786 nucleosome, GO:0003677 DNA binding, GO:0005634 nucleus, GO:0006334 nucleosome assembly	Histone H4
Phyca11_547376	GO:0019787 ubiquitin-like protein transferase activity, GO:0043687 post-translational protein modification, GO:0051246 regulation of protein metabolic process ,	Ubiquitin-protein ligase
Phyca11_547562	GO:0019787 ubiquitin-like protein transferase activity, GO:0043687 post-translational protein modification, GO:0051246 regulation of protein metabolic process ,	NADH:flavin oxidoreductase/12-oxophytodienoate reductase
Phyca11_563203	GO:0003824 catalytic activity, GO:0005488 binding, GO:0008152 metabolic process, GO:0008270 zinc ion binding, GO:0016491 oxidoreductase activity, GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, GO:0048037 cofactor binding ,	Alcohol dehydrogenase, class V
Phyca11_566774	-	-
Phyca11_569873	GO:0000786 nucleosome, GO:0003677 DNA binding, GO:0005634 nucleus, GO:0006334 nucleosome assembly	Histone 2A
Phyca11_576734	GO:0003924 GTPase activity, GO:0005198 structural molecule activity, GO:0005525 GTP binding, GO:0005874 microtubule, GO:0007017 microtubule-based process, GO:0007018 microtubule-based movement, GO:0043234 protein complex, GO:0051258 protein polymerization	Beta tubulin
Phyca11_8318	GO:0003924 GTPase activity, GO:0005515 protein binding, GO:0005524 ATP binding, GO:0005525 GTP binding, GO:0005622 intracellular, GO:0006355 regulation of transcription, DNA-templated, GO:0006886 intracellular protein transport, GO:0006913 nucleocytoplasmic transport, GO:0007165 signal transduction, GO:0007264 small GTPase mediated signal transduction, GO:0008134 transcription factor binding, GO:0015031 protein transport	GTPase Rab1/YPT1, small G protein superfamily, and related GTP-binding proteins
Phyca11_96761	GO:0006457 protein folding	FKBP-type peptidyl-prolyl cis-trans isomerase
Phyca11_97171	GO:0006118 obsolete electron transport, GO:0009055 electron carrier activity, GO:0016491 oxidoreductase activity, GO:0016651 oxidoreductase activity, acting on NAD(P)H, GO:0051536 iron-sulfur cluster binding ,	NADH:ubiquinone oxidoreductase, NDUFS8/23 kDa subunit
Induced in GC		
Phyca11_108059	GO:0003676 nucleic acid binding, GO:0004518 nuclease activity, GO:0005515 protein binding, GO:0016246 RNA interference, GO:0016442 RISC complex ,	Transcriptional coactivator p100
Phyca11_109772	GO:0003824 catalytic activity, GO:0016491 oxidoreductase activity, GO:0050660 flavin adenine dinucleotide binding ,	Proteins containing the FAD binding domain
Phyca11_111740	GO:0005524 ATP binding	Molecular chaperones HSP105/HSP110/SSE1, HSP70 superfamily
Phyca11_116241	GO:0000166 nucleotide binding, GO:0005524 ATP binding, GO:0005634 nucleus, GO:0005737 cytoplasm, GO:0016787 hydrolase activity, GO:0017111 nucleoside-triphosphatase activity, GO:0030163 protein catabolic process ,	26S proteasome regulatory complex, ATPase RPT4

Phyca11_116917	GO:0004197 cysteine-type endopeptidase activity, GO:0006508 proteolysis, GO:0008234 cysteine-type peptidase activity ,	Cysteine proteinase Cathepsin L
Phyca11_119319	GO:0003861 3-isopropylmalate dehydratase activity, GO:0008152 metabolic process, GO:0009098 leucine biosynthetic process, GO:0009316 3-isopropylmalate dehydratase complex	Aconitase/homoaconitase (aconitase superfamily)
Phyca11_120320	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004823 leucine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006429 leucyl-tRNA aminoacylation	Leucyl-tRNA synthetase
Phyca11_121219	GO:0003824 catalytic activity, GO:0004637 phosphoribosylamine-glycine ligase activity, GO:0004641 phosphoribosylformylglycinamidine cyclo-ligase activity, GO:0004644 phosphoribosylglycinamide formyltransferase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006188 IMP biosynthetic process, GO:0006189 'de novo' IMP biosynthetic process, GO:0008168 methyltransferase activity, GO:0009058 biosynthetic process, GO:0009113 purine nucleobase biosynthetic process, GO:0016742 hydroxymethyl-, formyl- and related transferase activity	Glycinamide ribonucleotide synthetase (GARS)/Aminoimidazole ribonucleotide synthetase (AIRS)
Phyca11_122536	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004831 tyrosine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006437 tyrosyl-tRNA aminoacylation	Tyrosyl-tRNA synthetase, cytoplasmic
Phyca11_132044	GO:0004585 ornithine carbamoyltransferase activity, GO:0006520 cellular amino acid metabolic process, GO:0009348 ornithine carbamoyltransferase complex, GO:0016597 amino acid binding, GO:0016743 carboxyl- or carbamoyltransferase activity ,	Ornithine carbamoyltransferase OTC/ARG3
Phyca11_13754	-	WD40 repeat stress protein/actin interacting protein
Phyca11_17928	GO:0004351 glutamate decarboxylase activity, GO:0006536 glutamate metabolic process, GO:0016831 carboxy-lyase activity, GO:0019752 carboxylic acid metabolic process, GO:0030170 pyridoxal phosphate binding ,	Glutamate decarboxylase/sphingosine phosphate lyase
Phyca11_18538	GO:0008483 transaminase activity, GO:0030170 pyridoxal phosphate binding	Alanine-glyoxylate aminotransferase AGT2
Phyca11_21714	GO:0004175 endopeptidase activity, GO:0004298 threonine-type endopeptidase activity, GO:0005839 proteasome core complex, GO:0006511 ubiquitin-dependent protein catabolic process	20S proteasome, regulatory subunit alpha type PSMA1/PRE5
Phyca11_36446	GO:0006525 arginine metabolic process, GO:0008483 transaminase activity, GO:0030170 pyridoxal phosphate binding ,	Acetylornithine aminotransferase
Phyca11_502585	GO:0009058 biosynthetic process, GO:0016769 transferase activity, transferring nitrogenous groups, GO:0030170 pyridoxal phosphate binding ,	Aromatic amino acid aminotransferase and related proteins
Phyca11_502885	GO:0003824 catalytic activity, GO:0004615 phosphomannomutase activity, GO:0005737 cytoplasm, GO:0008152 metabolic process, GO:0019307 mannose biosynthetic process ,	Phosphomannomutase
Phyca11_502951	-	Glutathione S-transferase
Phyca11_503099	GO:0009306 protein secretion, GO:0015450 P-P-bond-hydrolysis-driven protein transmembrane transporter activity, GO:0016020 membrane ,	Transport protein Sec61, alpha subunit
Phyca11_503447	GO:0005515 protein binding, GO:0005524 ATP binding, GO:0006457 protein folding, GO:0044267 cellular protein metabolic process, GO:0051082 unfolded protein binding ,	Chaperonin complex component, TCP-1 beta subunit (CCT2)
Phyca11_503535	GO:0003824 catalytic activity, GO:0003852 2-isopropylmalate synthase activity, GO:0009098 leucine biosynthetic process, GO:0019752 carboxylic acid metabolic process, GO:0046912 transferase activity, transferring acyl groups, acyl groups converted into alkyl on transfer ,	Alpha-isopropylmalate synthase/homocitrate synthase
Phyca11_503562	GO:0000166 nucleotide binding, GO:0005488 binding, GO:0005524 ATP binding, GO:0005634 nucleus, GO:0016887 ATPase activity, GO:0017111 nucleoside-triphosphatase activity	ATPase component of ABC transporters with duplicated ATPase domains/Translation elongation factor EF-3b
Phyca11_503635	GO:0009058 biosynthetic process, GO:0016769 transferase activity, transferring nitrogenous groups, GO:0030170 pyridoxal phosphate binding ,	Alanine aminotransferase
Phyca11_503639	GO:0003872 6-phosphofructokinase activity, GO:0005945 6-phosphofructokinase complex, GO:0006096 glycolytic process ,	Pyrophosphate-dependent phosphofructo-1-kinase

Phyca11_503977	GO:0009058 biosynthetic process, GO:0016769 transferase activity, transferring nitrogenous groups, GO:0030170 pyridoxal phosphate binding ,	-
Phyca11_503984	GO:0003676 nucleic acid binding, GO:0008270 zinc ion binding	Predicted RNA-binding protein containing PIN domain and invovled in translation or RNA processing
Phyca11_504193	GO:0003676 nucleic acid binding, GO:0003824 catalytic activity, GO:0004642 phosphoribosylformylglycinamidine synthase activity, GO:0005622 intracellular, GO:0006189 'de novo' IMP biosynthetic process, GO:0008270 zinc ion binding	Phosphoribosylformylglycinamidine synthase
Phyca11_504262	GO:0000166 nucleotide binding, GO:0005488 binding, GO:0005524 ATP binding, GO:0016787 hydrolase activity, GO:0017111 nucleoside-triphosphatase activity ,	AAA+-type ATPase
Phyca11_504517	GO:0005737 cytoplasm, GO:0006118 obsolete electron transport, GO:0016491 oxidoreductase activity, GO:0045454 cell redox homeostasis, GO:0050660 flavin adenine dinucleotide binding ,	Pyridine nucleotide-disulphide oxidoreductase
Phyca11_504831	GO:0003824 catalytic activity, GO:0004412 homoserine dehydrogenase activity, GO:0005488 binding, GO:0008152 metabolic process, GO:0008652 cellular amino acid biosynthetic process ,	Homoserine dehydrogenase
Phyca11_504894	GO:0003824 catalytic activity, GO:0003878 ATP citrate synthase activity, GO:0004775 succinate-CoA ligase (ADP-forming) activity, GO:0005488 binding, GO:0005737 cytoplasm, GO:0008152 metabolic process, GO:0044262 cellular carbohydrate metabolic process, GO:0046912 transferase activity, transferring acyl groups, acyl groups converted into alkyl on transfer	ATP-citrate lyase
Phyca11_505046	GO:0003676 nucleic acid binding, GO:0003723 RNA binding, GO:0003743 translation initiation factor activity, GO:0006413 translational initiation	Translation initiation factor 3, subunit b (eIF-3b)
Phyca11_505109	-	-
Phyca11_505794	GO:0004298 threonine-type endopeptidase activity, GO:0005839 proteasome core complex, GO:0006511 ubiquitin-dependent protein catabolic process ,	20S proteasome, regulatory subunit alpha type PSMA2/PRE8
Phyca11_506362	GO:0003824 catalytic activity, GO:0005488 binding, GO:0008152 metabolic process, GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, GO:0048037 cofactor binding, GO:0051287 NAD binding	D-3-phosphoglycerate dehydrogenase, D-isomer-specific 2-hydroxy acid dehydrogenase superfamily
Phyca11_506568	GO:0006457 protein folding, GO:0031072 heat shock protein binding, GO:0051082 unfolded protein binding ,	Molecular chaperone (DnaJ superfamily)
Phyca11_506596	GO:0004066 asparagine synthase (glutamine-hydrolyzing) activity, GO:0006529 asparagine biosynthetic process, GO:0008152 metabolic process ,	Asparagine synthase (glutamine-hydrolyzing)
Phyca11_506857	GO:0003824 catalytic activity, GO:0004802 transketolase activity, GO:0008152 metabolic process ,	Transketolase
Phyca11_507091	GO:0003779 actin binding, GO:0005488 binding, GO:0007010 cytoskeleton organization ,	Adenylate cyclase-associated protein (CAP/Srv2p)
Phyca11_507344	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004828 serine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006434 seryl-tRNA aminoacylation	Seryl-tRNA synthetase
Phyca11_507390	GO:0003984 acetolactate synthase activity, GO:0008152 metabolic process, GO:0009082 branched-chain amino acid biosynthetic process, GO:0016597 amino acid binding	Acetolactate synthase, small subunit
Phyca11_507479	GO:0008152 metabolic process, GO:0016624 oxidoreductase activity, acting on the aldehyde or oxo group of donors, disulfide as acceptor	Pyruvate dehydrogenase E1, alpha subunit
Phyca11_507542	GO:0016209 antioxidant activity, GO:0016491 oxidoreductase activity, GO:0045454 cell redox homeostasis ,	Alkyl hydroperoxide reductase, thiol specific antioxidant and related enzymes
Phyca11_507640	GO:0003824 catalytic activity, GO:0004455 ketol-acid reductoisomerase activity, GO:0005488 binding, GO:0008152 metabolic process, GO:0009082 branched-chain amino acid biosynthetic process, GO:0016491 oxidoreductase activity	-
Phyca11_508386	GO:0005093 Rab GDP-dissociation inhibitor activity, GO:0015031 protein transport, GO:0043087 regulation of GTPase activity ,	RAB proteins geranylgeranyltransferase component A (RAB escort protein)
Phyca11_508572	GO:0000166 nucleotide binding, GO:0003676 nucleic acid binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004824 lysine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006430 lysyl-tRNA aminoacylation ,	Lysyl-tRNA synthetase (class II)

Phyca11_508653	GO:0003824 catalytic activity, GO:0003987 acetate-CoA ligase activity, GO:0008152 metabolic process, GO:0016208 AMP binding	Acyl-CoA synthetase
Phyca11_508846	GO:0004372 glycine hydroxymethyltransferase activity, GO:0006544 glycine metabolic process, GO:0006563 L-serine metabolic process ,	Glycine/serine hydroxymethyltransferase
Phyca11_509055	GO:0003779 actin binding, GO:0007010 cytoskeleton organization	Actin regulatory proteins (gelsolin/villin family)
Phyca11_509385	GO:0000287 magnesium ion binding, GO:0004427 inorganic diphosphatase activity, GO:0005509 calcium ion binding, GO:0005737 cytoplasm, GO:0006796 phosphate-containing compound metabolic process ,	Inorganic pyrophosphatase/Nucleosome remodeling factor, subunit NURF38
Phyca11_509479	GO:0003995 acyl-CoA dehydrogenase activity, GO:0006118 obsolete electron transport, GO:0008152 metabolic process, GO:0016627 oxidoreductase activity, acting on the CH-CH group of donors	Medium-chain acyl-CoA dehydrogenase
Phyca11_509552	GO:0004356 glutamate-ammonia ligase activity, GO:0006542 glutamine biosynthetic process, GO:0006807 nitrogen compound metabolic process ,	Glutamine synthetase
Phyca11_509625	GO:0003824 catalytic activity, GO:0004345 glucose-6-phosphate dehydrogenase activity, GO:0005488 binding, GO:0006006 glucose metabolic process, GO:0008152 metabolic process ,	Glucose-6-phosphate 1-dehydrogenase
Phyca11_509634	GO:0003824 catalytic activity, GO:0004617 phosphoglycerate dehydrogenase activity, GO:0005215 transporter activity, GO:0005488 binding, GO:0006564 L-serine biosynthetic process, GO:0006810 transport, GO:0008152 metabolic process, GO:0009058 biosynthetic process, GO:0016020 membrane, GO:0016597 amino acid binding, GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, GO:0048037 cofactor binding, GO:0051287 NAD binding,	D-3-phosphoglycerate dehydrogenase, D-isomer-specific 2-hydroxy acid dehydrogenase superfamily
Phyca11_509774	GO:0003871 5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase activity, GO:0008270 zinc ion binding, GO:0008652 cellular amino acid biosynthetic process, GO:0009086 methionine biosynthetic process	Methionine synthase II (cobalamin-independent)
Phyca11_510165	GO:0000287 magnesium ion binding, GO:0003824 catalytic activity, GO:0004743 pyruvate kinase activity, GO:0006096 glycolytic process, GO:0030955 potassium ion binding ,	Pyruvate kinase
Phyca11_510275	GO:0003824 catalytic activity, GO:0004056 argininosuccinate lyase activity, GO:0042450 arginine biosynthetic process via ornithine ,	Argininosuccinate lyase
Phyca11_510471	-	Translation initiation factor 3, subunit i (eIF-3i)/TGF-beta receptor-interacting protein (TRIP-1)
Phyca11_510618	GO:0003824 catalytic activity, GO:0005524 ATP binding, GO:0006810 transport, GO:0006812 cation transport, GO:0008152 metabolic process, GO:0015662 ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism, GO:0015992 proton transport, GO:0016020 membrane, GO:0016021 integral component of membrane, GO:0016820 hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances, GO:0016887 ATPase activity ,	Plasma membrane H+-transporting ATPase
Phyca11_510725	GO:0003824 catalytic activity, GO:0008152 metabolic process	Stationary phase-induced protein, SOR/SNZ family
Phyca11_510858	GO:0004368 glycerol-3-phosphate dehydrogenase activity, GO:0006072 glycerol-3-phosphate metabolic process, GO:0009331 glycerol-3-phosphate dehydrogenase complex, GO:0016491 oxidoreductase activity	Glycerol-3-phosphate dehydrogenase
Phyca11_511077	GO:0000287 magnesium ion binding, GO:0003824 catalytic activity, GO:0004743 pyruvate kinase activity, GO:0006096 glycolytic process, GO:0030955 potassium ion binding ,	Pyruvate kinase
Phyca11_511132	GO:0005634 nucleus, GO:0006334 nucleosome assembly	Nucleosome assembly protein NAP-1
Phyca11_511318	GO:0004348 glucosylceramidase activity, GO:0005764 lysosome, GO:0006665 sphingolipid metabolic process, GO:0007040 lysosome organization	Beta-glucocerebrosidase
Phyca11_511385	GO:0005975 carbohydrate metabolic process, GO:0008152 metabolic process, GO:0016779 nucleotidyltransferase activity, GO:0016868 intramolecular transferase activity, phosphotransferases	Phosphoglucomutase
Phyca11_511466	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004829 threonine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006435 threonyl-tRNA aminoacylation,	Threonyl-tRNA synthetase

	GO:0016876 ligase activity, forming aminoacyl-tRNA and related compounds, GO:0043039 tRNA aminoacylation	
Phyca11_511519	GO:0003746 translation elongation factor activity, GO:0003924 GTPase activity, GO:0005525 GTP binding, GO:0005622 intracellular, GO:0006414 translational elongation ,	Mitochondrial translation elongation factor Tu
Phyca11_511707	GO:0003735 structural constituent of ribosome, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006412 translation	40S ribosomal protein S12
Phyca11_511793	GO:0003824 catalytic activity, GO:0005488 binding, GO:0008152 metabolic process, GO:0008270 zinc ion binding, GO:0016491 oxidoreductase activity ,	Alcohol dehydrogenase, class V
Phyca11_511870	GO:0005741 mitochondrial outer membrane, GO:0006820 anion transport, GO:0008308 voltage-gated anion channel activity ,	Porin/voltage-dependent anion-selective channel protein
Phyca11_511879	GO:0005737 cytoplasm	Microtubule-binding protein (translationally controlled tumor protein)
Phyca11_512063	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004819 glutamine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006424 glutamyl-tRNA aminoacylation, GO:0006425 glutaminyl-tRNA aminoacylation ,	Glutaminyl-tRNA synthetase
Phyca11_525668	GO:0003824 catalytic activity, GO:0005488 binding, GO:0006118 obsolete electron transport, GO:0008152 metabolic process, GO:0016491 oxidoreductase activity, GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, GO:0051287 NAD binding ,	UDP-glucose/GDP-mannose dehydrogenase
Phyca11_526928	GO:0003824 catalytic activity, GO:0004086 obsolete carbamoyl-phosphate synthase activity, GO:0005524 ATP binding, GO:0006541 glutamine metabolic process, GO:0006807 nitrogen compound metabolic process, GO:0008152 metabolic process	Multifunctional pyrimidine synthesis protein CAD (includes carbamoyl-phophate synthetase, aspartate transcarbamylase, and glutamine amidotransferase)
Phyca11_526964	GO:0003735 structural constituent of ribosome, GO:0003824 catalytic activity, GO:0004735 pyrroline-5-carboxylate reductase activity, GO:0005488 binding, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006118 obsolete electron transport, GO:0006412 translation, GO:0006561 proline biosynthetic process, GO:0008152 metabolic process	Pyrroline-5-carboxylate reductase
Phyca11_527152	GO:0016746 transferase activity, transferring acyl groups	Carnitine O-acyltransferase CPT2/YAT1
Phyca11_529555	GO:0003995 acyl-CoA dehydrogenase activity, GO:0006118 obsolete electron transport, GO:0008152 metabolic process, GO:0016627 oxidoreductase activity, acting on the CH-CH group of donors	Very-long-chain acyl-CoA dehydrogenase
Phyca11_529861	GO:0003824 catalytic activity, GO:0005488 binding, GO:0006520 cellular amino acid metabolic process, GO:0008152 metabolic process, GO:0016491 oxidoreductase activity ,	Glutamate/leucine/phenylalanine/valine dehydrogenases
Phyca11_529913	GO:0005515 protein binding, GO:0005524 ATP binding, GO:0006457 protein folding, GO:0044267 cellular protein metabolic process, GO:0051082 unfolded protein binding ,	Chaperonin complex component, TCP-1 delta subunit (CCT4)
Phyca11_530278	GO:0003824 catalytic activity, GO:0005488 binding, GO:0006631 fatty acid metabolic process, GO:0008152 metabolic process, GO:0016491 oxidoreductase activity ,	3-hydroxyacyl-CoA dehydrogenase
Phyca11_530893	GO:0004096 catalase activity, GO:0004601 peroxidase activity, GO:0006118 obsolete electron transport, GO:0006979 response to oxidative stress, GO:0020037 heme binding ,	-
Phyca11_532200	GO:0000059 protein import into nucleus, docking, GO:0005488 binding, GO:0005634 nucleus, GO:0005643 nuclear pore, GO:0005737 cytoplasm, GO:0006886 intracellular protein transport, GO:0008565 protein transporter activity ,	Karyopherin (importin) beta 3
Phyca11_532207	GO:0003676 nucleic acid binding, GO:0003677 DNA binding, GO:0005524 ATP binding, GO:0005622 intracellular, GO:0008270 zinc ion binding ,	Molecular chaperones HSP70/HSC70, HSP70 superfamily
Phyca11_533182	GO:0003824 catalytic activity, GO:0005488 binding, GO:0006520 cellular amino acid metabolic process, GO:0008152 metabolic process, GO:0016491 oxidoreductase activity ,	Glutamate/leucine/phenylalanine/valine dehydrogenases
Phyca11_533830	GO:0004055 argininosuccinate synthase activity, GO:0005524 ATP binding, GO:0006526 arginine biosynthetic process ,	Argininosuccinate synthase
Phyca11_534256	GO:0005515 protein binding, GO:0008152 metabolic process, GO:0016746 transferase activity, transferring acyl groups ,	Dihydrolipoamide acetyltransferase

Phyca11_534354	GO:0003677 DNA binding, GO:0003824 catalytic activity, GO:0008152 metabolic process, GO:0046983 protein dimerization activity	Acetyl-CoA acetyltransferase
Phyca11_534535	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004822 isoleucine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006428 isoleucyl-tRNA aminoacylation	Isoleucyl-tRNA synthetase
Phyca11_536029	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004827 proline-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006433 prolyl-tRNA aminoacylation	Prolyl-tRNA synthetase
Phyca11_536632	GO:0003824 catalytic activity, GO:0003922 GMP synthase (glutamine-hydrolyzing) activity, GO:0004672 protein kinase activity, GO:0004674 protein serine/threonine kinase activity, GO:0004713 protein tyrosine kinase activity, GO:0004808 tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006164 purine nucleotide biosynthetic process, GO:0006177 GMP biosynthetic process, GO:0006468 protein phosphorylation, GO:0006541 glutamine metabolic process, GO:0008033 tRNA processing, GO:0008152 metabolic process	GMP synthase
Phyca11_5370	GO:0005215 transporter activity, GO:0005488 binding, GO:0005509 calcium ion binding, GO:0005743 mitochondrial inner membrane, GO:0006810 transport, GO:0016020 membrane	Mitochondrial tricarboxylate/dicarboxylate carrier proteins
Phyca11_537682	GO:0003676 nucleic acid binding, GO:0005622 intracellular, GO:0006810 transport ,	RasGAP SH3 binding protein rasputin, contains NTF2 and RRM domains
Phyca11_537937	GO:0004252 serine-type endopeptidase activity, GO:0004287 obsolete prolyl oligopeptidase activity, GO:0006508 proteolysis, GO:0008236 serine-type peptidase activity	Predicted serine protease
Phyca11_538407	GO:0003824 catalytic activity, GO:0006520 cellular amino acid metabolic process, GO:0008483 transaminase activity, GO:0009058 biosynthetic process, GO:0016769 transferase activity, transferring nitrogenous groups, GO:0030170 pyridoxal phosphate binding	Aspartate aminotransferase/Glutamic oxaloacetic transaminase AAT1/GOT2
Phyca11_538943	GO:0003824 catalytic activity, GO:0004086 obsolete carbamoyl-phosphate synthase activity, GO:0005524 ATP binding, GO:0006541 glutamine metabolic process, GO:0006807 nitrogen compound metabolic process, GO:0008152 metabolic process	Multifunctional pyrimidine synthesis protein CAD (includes carbamoyl-phophate synthetase, aspartate transcarbamylase, and glutamine amidotransferase)
Phyca11_539198	GO:0006118 obsolete electron transport, GO:0006120 mitochondrial electron transport, NADH to ubiquinone, GO:0008137 NADH dehydrogenase (ubiquinone) activity, GO:0010181 FMN binding, GO:0016651 oxidoreductase activity, acting on NAD(P)H, GO:0051287 NAD binding, GO:0051539 4 iron, 4 sulfur cluster binding ,	NADH:ubiquinone oxidoreductase, NDUFV1/51kDa subunit
Phyca11_5400	-	Molecular co-chaperone STI1
Phyca11_542172	GO:0005975 carbohydrate metabolic process, GO:0042578 phosphoric ester hydrolase activity	Fructose-1,6-bisphosphatase
Phyca11_543889	GO:0003824 catalytic activity	Spermidine synthase
Phyca11_544083	-	Molecular co-chaperone STI1
Phyca11_545598	GO:0003676 nucleic acid binding, GO:0004813 alanine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006419 alanyl-tRNA aminoacylation, GO:0016876 ligase activity, forming aminoacyl-tRNA and related compounds, GO:0043039 tRNA aminoacylation	Alanyl-tRNA synthetase
Phyca11_548602	GO:0003676 nucleic acid binding, GO:0004386 helicase activity, GO:0005524 ATP binding, GO:0008026 ATP-dependent helicase activity	ATP-dependent RNA helicase
Phyca11_548878	GO:0000166 nucleotide binding, GO:0003723 RNA binding, GO:0004826 phenylalanine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006432 phenylalanyl-tRNA aminoacylation ,	Phenylalanyl-tRNA synthetase beta subunit
Phyca11_549506	GO:0003676 nucleic acid binding, GO:0003723 RNA binding, GO:0016071 mRNA metabolic process ,	Polyadenylate-binding protein (RRM superfamily)
Phyca11_549899	GO:0003824 catalytic activity, GO:0004735 pyrroline-5-carboxylate reductase activity, GO:0005488 binding, GO:0006118 obsolete electron transport, GO:0006561 proline biosynthetic process, GO:0008152 metabolic process	-

Phyca11_551783	-	Histone acetyltransferase SAGA/ADA, catalytic subunit PCAF/GCN5 and related proteins
Phyca11_551916	-	-
Phyca11_553648	GO:0000287 magnesium ion binding, GO:0003824 catalytic activity, GO:0004743 pyruvate kinase activity, GO:0006096 glycolytic process, GO:0030955 potassium ion binding ,	Pyruvate kinase
Phyca11_554319	GO:0004298 threonine-type endopeptidase activity, GO:0005839 proteasome core complex, GO:0006511 ubiquitin-dependent protein catabolic process ,	20S proteasome, regulatory subunit beta type PSMB2/PRE1
Phyca11_556236	GO:0008483 transaminase activity, GO:0030170 pyridoxal phosphate binding	Ornithine aminotransferase
Phyca11_556289	GO:0004047 aminomethyltransferase activity, GO:0005737 cytoplasm, GO:0006546 glycine catabolic process ,	Aminomethyl transferase
Phyca11_558677	GO:0005488 binding, GO:0005515 protein binding, GO:0006461 protein complex assembly, GO:0006886 intracellular protein transport, GO:0008565 protein transporter activity, GO:0016192 vesicle-mediated transport, GO:0030117 membrane coat, GO:0030131 clathrin adaptor complex	Vesicle coat complex AP-2, alpha subunit
Phyca11_559116	GO:0003994 aconitate hydratase activity, GO:0006099 tricarboxylic acid cycle, GO:0008152 metabolic process, GO:0051539 4 iron, 4 sulfur cluster binding	Aconitase/homoaconitase (aconitase superfamily)
Phyca11_562095	GO:0005198 structural molecule activity, GO:0005488 binding, GO:0005515 protein binding, GO:0006461 protein complex assembly, GO:0006886 intracellular protein transport, GO:0008565 protein transporter activity, GO:0016192 vesicle-mediated transport, GO:0030130 clathrin coat of trans-Golgi network vesicle, GO:0030132 clathrin coat of coated pit ,	Vesicle coat protein clathrin, heavy chain
Phyca11_562442	GO:0008152 metabolic process, GO:0016491 oxidoreductase activity	Aldehyde dehydrogenase
Phyca11_563206	GO:0003824 catalytic activity, GO:0005524 ATP binding, GO:0008152 metabolic process, GO:0009374 biotin binding, GO:0016874 ligase activity ,	3-Methylcrotonyl-CoA carboxylase, biotin-containing subunit/Propionyl-CoA carboxylase, alpha chain/Acetyl-CoA carboxylase, biotin carboxylase subunit
Phyca11_564009	GO:0004591 oxoglutarate dehydrogenase (succinyl-transferring) activity, GO:0006096 glycolytic process, GO:0008152 metabolic process, GO:0016624 oxidoreductase activity, acting on the aldehyde or oxo group of donors, disulfide as acceptor, GO:0030976 thiamine pyrophosphate binding ,	2-oxoglutarate dehydrogenase, E1 subunit
Phyca11_567952	GO:0005525 GTP binding, GO:0005622 intracellular	Predicted GTP-binding protein (ODN superfamily)
Phyca11_569273	GO:0005515 protein binding, GO:0005524 ATP binding, GO:0006457 protein folding, GO:0044267 cellular protein metabolic process, GO:0051082 unfolded protein binding ,	Chaperonin complex component, TCP-1 alpha subunit (CCT1)
Phyca11_569541	GO:0003824 catalytic activity, GO:0004664 prephenate dehydratase activity, GO:0009058 biosynthetic process, GO:0009094 L-phenylalanine biosynthetic process, GO:0016769 transferase activity, transferring nitrogenous groups, GO:0016847 1-aminocyclopropane-1-carboxylate synthase activity, GO:0030170 pyridoxal phosphate binding ,	Kynurenine aminotransferase, glutamine transaminase K
Phyca11_569720	GO:0005515 protein binding, GO:0005524 ATP binding, GO:0006457 protein folding, GO:0044267 cellular protein metabolic process, GO:0051082 unfolded protein binding ,	Chaperonin complex component, TCP-1 gamma subunit (CCT3)
Phyca11_573119	GO:0000221 vacuolar proton-transporting V-type ATPase, V1 domain, GO:0000300 obsolete peripheral to membrane of membrane fraction, GO:0005488 binding, GO:0005524 ATP binding, GO:0046961 proton-transporting ATPase activity, rotational mechanism ,	Vacuolar H+-ATPase V1 sector, subunit H
Phyca11_573714	-	Uncharacterized conserved protein
Phyca11_573963	GO:0008152 metabolic process, GO:0016491 oxidoreductase activity	Aldehyde dehydrogenase
Phyca11_574456	GO:0016301 kinase activity, GO:0016772 transferase activity, transferring phosphorus-containing groups	Creatine kinases
Phyca11_574695	GO:0003824 catalytic activity, GO:0003937 IMP cyclohydrolase activity, GO:0004643 phosphoribosylaminoimidazolecarboxamide formyltransferase activity, GO:0006164 purine nucleotide biosynthetic process, GO:0006188 IMP biosynthetic process ,	AICAR transformylase/IMP cyclohydrolase/methylglyoxal synthase
Phyca11_575325	GO:0004491 methylmalonate-semialdehyde dehydrogenase (acylating) activity, GO:0006573 valine metabolic process, GO:0008152 metabolic process, GO:0016491 oxidoreductase activity	Methylmalonate semialdehyde dehydrogenase

Phyca11_577084	GO:0003924 GTPase activity, GO:0005525 GTP binding	Polypeptide release factor 3
Phyca11_577399	GO:0000049 tRNA binding, GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004825 methionine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006431 methionyl-tRNA aminoacylation ,	Methionyl-tRNA synthetase
Phyca11_577453	-	26S proteasome regulatory complex, subunit RPN3/PSMD3
Phyca11_5784	GO:0003824 catalytic activity, GO:0008152 metabolic process	Acetyl-CoA acetyltransferase
Phyca11_6173	GO:0005488 binding	26S proteasome regulatory complex, subunit RPN2/PSMD1
Phyca11_8284	GO:0003824 catalytic activity, GO:0004474 malate synthase activity, GO:0006097 glyoxylate cycle ,	Malate synthase
Phyca11_8634	-	Dihydrolipoamide acetyltransferase
Phyca11_8926	GO:0005515 protein binding, GO:0005524 ATP binding, GO:0006457 protein folding, GO:0044267 cellular protein metabolic process, GO:0051082 unfolded protein binding ,	Chaperonin complex component, TCP-1 theta subunit (CCT8)
Phyca11_96835	-	26S proteasome regulatory complex, subunit RPN5/PSMD12
Phyca11_98041	GO:0005488 binding	26S proteasome regulatory complex, subunit RPN1/PSMD2
Phyca11_98712	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004818 glutamate-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006424 glutamyl-tRNA aminoacylation	Glutamyl-tRNA synthetase
Phyca11_99958	GO:0003676 nucleic acid binding, GO:0004386 helicase activity, GO:0005524 ATP binding, GO:0008026 ATP-dependent helicase activity	ATP-dependent RNA helicase

Supplementary Table 4: GO terms and KOG annotations for all DE proteins from TE versus CE comparisons

Phyca 11 number	Gene Ontology	KOG
Induced in T		
Phyca11_101373	GO:0003735 structural constituent of ribosome, GO:0005622 intracellular, GO:0005840 ribosome, GO:0006412 translation	40S ribosomal protein S21
Phyca11_110703	GO:0004553 hydrolase activity, hydrolyzing O-glycosyl compounds, GO:0005975 carbohydrate metabolic process	Beta-glucosidase, lactase phlorizinhydrolase, and related proteins
Phyca11_120320	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004823 leucine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006429 leucyl-tRNA aminoacylation	Leucyl-tRNA synthetase
Phyca11_503721	-	-
Phyca11_504650	GO:0003924 GTPase activity, GO:0005515 protein binding, GO:0005524 ATP binding, GO:0005525 GTP binding, GO:0005622 intracellular, GO:0006355 regulation of transcription, DNA-templated, GO:0006886 intracellular protein transport, GO:0006913 nucleocytoplasmic transport, GO:0007165 signal transduction, GO:0007264 small GTPase mediated signal transduction, GO:0008134 transcription factor binding, GO:0015031 protein transport	GTPase Rab2, small G protein superfamily
Phyca11_505974	GO:0004427 inorganic diphosphatase activity, GO:0009678 hydrogen-translocating pyrophosphatase activity, GO:0015992 proton transport, GO:0016020 membrane	-
Phyca11_508140	-	-
Phyca11_508411	-	26S proteasome regulatory complex, subunit PSMD10
Phyca11_510755	GO:0000166 nucleotide binding, GO:0005515 protein binding, GO:0005524 ATP binding, GO:0017111 nucleoside-triphosphatase activity, GO:0019538 protein metabolic process ,	Chaperone HSP104 and related ATP-dependent Clp proteases
Phyca11_534354	GO:0003677 DNA binding, GO:0003824 catalytic activity, GO:0008152 metabolic process, GO:0046983 protein dimerization activity	Acetyl-CoA acetyltransferase
Phyca11_538207	GO:0000166 nucleotide binding, GO:0003676 nucleic acid binding, GO:0004386 helicase activity, GO:0005524 ATP binding, GO:0008026 ATP-dependent helicase activity, GO:0017111 nucleoside-triphosphatase activity	DEAH-box RNA helicase
Induced in C		
Phyca11_111740	GO:0005524 ATP binding	Molecular chaperones HSP105/HSP110/SSE1, HSP70 superfamily
Phyca11_118794	GO:0004089 carbonate dehydratase activity, GO:0008270 zinc ion binding, GO:0015976 carbon utilization	Predicted carbonic anhydrase involved in protection against oxidative damage
Phyca11_21714	GO:0004175 endopeptidase activity, GO:0004298 threonine-type endopeptidase activity, GO:0005839 proteasome core complex, GO:0006511 ubiquitin-dependent protein catabolic process	20S proteasome, regulatory subunit alpha type PSMA1/PRE5
Phyca11_36446	GO:0006525 arginine metabolic process, GO:0008483 transaminase activity, GO:0030170 pyridoxal phosphate binding	Acetylornithine aminotransferase
Phyca11_43072	GO:0003746 translation elongation factor activity, GO:0005853 eukaryotic translation elongation factor 1 complex, GO:0006414 translational elongation	Elongation factor 1 beta/delta chain
Phyca11_502951	-	Glutathione S-transferase
Phyca11_503562	GO:0000166 nucleotide binding, GO:0005488 binding, GO:0005524 ATP binding, GO:0005634 nucleus, GO:0016887 ATPase activity, GO:0017111 nucleoside-triphosphatase activity	ATPase component of ABC transporters with duplicated ATPase domains/Translation elongation factor EF-3b
Phyca11_503639	GO:0003872 6-phosphofructokinase activity, GO:0005945 6-phosphofructokinase complex, GO:0006096 glycolytic process	Pyrophosphate-dependent phosphofructo-1-kinase
Phyca11_504262	GO:0000166 nucleotide binding, GO:0005488 binding, GO:0005524 ATP binding, GO:0016787 hydrolase activity, GO:0017111 nucleoside-triphosphatase activity ,	AAA+-type ATPase

Phyca11_504839	GO:0006813 potassium ion transport, GO:0016021 integral component of membrane, GO:0016491 oxidoreductase activity	Voltage-gated shaker-like K+ channel, subunit beta/KCNAB
Phyca11_505083	GO:0003824 catalytic activity, GO:0008152 metabolic process	Long-chain acyl-CoA synthetases (AMP-forming)
Phyca11_505601	GO:0006118 obsolete electron transport, GO:0009055 electron carrier activity	Electron transfer flavoprotein, beta subunit
Phyca11_507344	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004828 serine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006434 seryl-tRNA aminoacylation	Seryl-tRNA synthetase
Phyca11_508846	GO:0004372 glycine hydroxymethyltransferase activity, GO:0006544 glycine metabolic process, GO:0006563 L-serine metabolic process	Glycine/serine hydroxymethyltransferase
Phyca11_509184	GO:0004177 aminopeptidase activity, GO:0005622 intracellular, GO:0006508 proteolysis	Predicted aminopeptidase of the M17 family
Phyca11_509200	GO:0003824 catalytic activity, GO:0005488 binding, GO:0008152 metabolic process	Predicted NAD-dependent oxidoreductase
Phyca11_509552	GO:0004356 glutamate-ammonia ligase activity, GO:0006542 glutamine biosynthetic process, GO:0006807 nitrogen compound metabolic process	Glutamine synthetase
Phyca11_509625	GO:0003824 catalytic activity, GO:0004345 glucose-6-phosphate dehydrogenase activity, GO:0005488 binding, GO:0006006 glucose metabolic process, GO:0008152 metabolic process ,	Glucose-6-phosphate 1-dehydrogenase
Phyca11_509774	GO:0003871 5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase activity, GO:0008270 zinc ion binding, GO:0008652 cellular amino acid biosynthetic process, GO:0009086 methionine biosynthetic process	Methionine synthase II (cobalamin-independent)
Phyca11_510042	GO:0004340 glucokinase activity, GO:0005524 ATP binding, GO:0006096 glycolytic process	-
Phyca11_511077	GO:0000287 magnesium ion binding, GO:0003824 catalytic activity, GO:0004743 pyruvate kinase activity, GO:0006096 glycolytic process, GO:0030955 potassium ion binding ,	Pyruvate kinase
Phyca11_511132	GO:0005634 nucleus, GO:0006334 nucleosome assembly	Nucleosome assembly protein NAP-1
Phyca11_511385	GO:0005975 carbohydrate metabolic process, GO:0008152 metabolic process, GO:0016779 nucleotidyltransferase activity, GO:0016868 intramolecular transferase activity, phosphotransferases	Phosphoglucomutase
Phyca11_511466	GO:0000166 nucleotide binding, GO:0004812 aminoacyl-tRNA ligase activity, GO:0004829 threonine-tRNA ligase activity, GO:0005524 ATP binding, GO:0005737 cytoplasm, GO:0006412 translation, GO:0006418 tRNA aminoacylation for protein translation, GO:0006435 threonyl-tRNA aminoacylation, GO:0016876 ligase activity, forming aminoacyl-tRNA and related compounds, GO:0043039 tRNA aminoacylation	Threonyl-tRNA synthetase
Phyca11_511870	GO:0005741 mitochondrial outer membrane, GO:0006820 anion transport, GO:0008308 voltage-gated anion channel activity	Porin/voltage-dependent anion-selective channel protein
Phyca11_525668	GO:0003824 catalytic activity, GO:0005488 binding, GO:0006118 obsolete electron transport, GO:0008152 metabolic process, GO:0016491 oxidoreductase activity, GO:0016616 oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor, GO:0051287 NAD binding ,	UDP-glucose/GDP-mannose dehydrogenase
Phyca11_533830	GO:0004055 argininosuccinate synthase activity, GO:0005524 ATP binding, GO:0006526 arginine biosynthetic process	Argininosuccinate synthase
Phyca11_538675	GO:0000786 nucleosome, GO:0003677 DNA binding, GO:0005634 nucleus, GO:0006334 nucleosome assembly	Histone H4
Phyca11_538943	GO:0003824 catalytic activity, GO:0004086 obsolete carbamoyl-phosphate synthase activity, GO:0005524 ATP binding, GO:0006541 glutamine metabolic process, GO:0006807 nitrogen compound metabolic process, GO:0008152 metabolic process	Multifunctional pyrimidine synthesis protein CAD (includes carbamoyl-phophate synthetase, aspartate transcarbamylase, and glutamine amidotransferase)
Phyca11_5400	-	Molecular co-chaperone STI1
Phyca11_542172	GO:0005975 carbohydrate metabolic process, GO:0042578 phosphoric ester hydrolase activity	Fructose-1,6-bisphosphatase
Phyca11_551325	GO:0003924 GTPase activity, GO:0005525 GTP binding	Translation initiation factor 2, gamma subunit (eIF-2gamma; GTPase)
Phyca11_553648	GO:0000287 magnesium ion binding, GO:0003824 catalytic activity, GO:0004743 pyruvate kinase activity, GO:0006096 glycolytic process, GO:0030955 potassium ion binding ,	Pyruvate kinase
Phyca11_559116	GO:0003994 aconitate hydratase activity, GO:0006099 tricarboxylic acid cycle, GO:0008152 metabolic process, GO:0051539 4 iron, 4 sulfur cluster binding	Aconitase/homoaconitase (aconitase superfamily)

Phyca11_562095	GO:0005198 structural molecule activity, GO:0005488 binding, GO:0005515 protein binding, GO:0006461 protein complex assembly, GO:0006886 intracellular protein transport, GO:0008565 protein transporter activity, GO:0016192 vesicle-mediated transport, GO:0030130 clathrin coat of trans-Golgi network vesicle, GO:0030132 clathrin coat of coated pit ,	Vesicle coat protein clathrin, heavy chain
Phyca11_562546	GO:0004174 electron-transferring-flavoprotein dehydrogenase activity, GO:0006118 obsolete electron transport	Electron transfer flavoprotein ubiquinone oxidoreductase
Phyca11_576877	GO:0003824 catalytic activity, GO:0005488 binding, GO:0008152 metabolic process, GO:0016491 oxidoreductase activity	Peroxisomal multifunctional beta-oxidation protein and related enzymes
Phyca11_8284	GO:0003824 catalytic activity, GO:0004474 malate synthase activity, GO:0006097 glyoxylate cycle	Malate synthase